Ambiguity without a State Space*

David S. Ahn†

This version: July 2007
First version: October 2003

Abstract

Many decisions involve both imprecise probabilities and intractable states of the world. Objective expected utility assumes unambiguous probabilities; subjective expected utility assumes a completely specified state space. This paper analyzes a third domain of preference: sets of consequential lotteries. Using this domain, we develop a theory of objective ambiguity without explicit reference to any state space. We characterize a representation that integrates a nonlinear transformation of first order expected utility with respect to a second order measure. The concavity of the transformation and the weighting of the measure capture ambiguity aversion. We propose a definition for comparative ambiguity aversion.

* *Review of Economic Studies* 71(1): 3–28, January 2008. This copy is posted with permission of The Review of Economic Studies Limited and Blackwell Publishing. The definitive version is available at www.blackwell-synergy.com.

† Department of Economics, U.C. Berkeley, 549 Evans Hall #3880, Berkeley, CA 94720; dahn@econ.berkeley.edu.
The author thanks Susan Athey, David Baron, Simon Board, Chris Chambers, Larry Epstein, Haluk Ergin, Yossi Feinberg, Peter Hammond, David Kreps, Mark Machina, Wojciech Olszewski, Andy Skrzypacz, Chris Tyson, and many seminar participants for comments; Bob Anderson, Oleh Nykyforchyn, Chris Shannon, Max Stintchombe, and especially Bill Zame for numerous technical clarifications of an earlier version of the paper; and Juuso Välimäki and two anonymous referees for suggesting helpful revisions. This paper is a revision of the second chapter of the author's dissertation, supervised by Bob Wilson, to whom he owes a special debt for encouragement and support.

1 Introduction

Consider a terminally ill patient whose doctor suggests two treatments. The first is an established pharmaceutical. Numerous published studies concur that this drug is successful in thirty percent of cases. The second is a new experimental surgery. Its preliminary trials suggest a success rate between twenty and forty-five percent. The two treatments are mutually exclusive, so the patient must choose between them. Can we help the patient by framing her problem with either the von Neumann–Morgenstern (1944) or Savage (1954) theory of choice under uncertainty?

We cannot frame the patient's problem in the standard von Neumann–Morgenstern (henceforth vNM) setting; the surgery is associated with an ambiguous range of possible success rates. This deficiency in the primitives, the tacit assumption of precise probability, cannot be salvaged by relaxing axioms. The subjective theory of Savage (1954) does not assume the probabilities of different outcomes are exogenously precise.[1] On the other hand, to invoke the Savage machinery, our patient must be able to: first, determine the relevant states of nature; second, decide how each choice assigns consequences to these states.

She fails both counts and cannot frame her problem using subjective utility theory. Regarding the state space, the patient has no medical training and does not understand what the relevant states are. She knows only the information presented by her physician, expressed entirely in the space of probabilities over consequences. Even given a comprehensive list of states, the designers of the experimental surgery are unsure which states would make the surgery more likely to be successful. Studies are needed exactly because the mapping from states to outcomes that actually represents this new medical procedure is still unknown. More generally, an inability to correctly formulate the state space or the acts is often the cause of the ambiguity in a decision problem. As with objective vNM theory, the deficiency is not any particular subjective utility representation, but is fundamental to the structure of an act.

There seem to be many situations where the decision maker cannot properly conceptualize the state space or the appropriate mappings, but still has to decide on a course of action. The primary methodological contribution of this paper is a formal language to discuss ambiguity without appealing to any state space, thus avoiding the technology of states and acts. The machinery of vNM is too simple to express the patient's problem; the machinery of Savage is too complicated. We study an alternative framework to analyze such decisions: sets of lotteries over consequences. Each set captures the possible distributions on consequences associated with a particular option. By enriching the domain of preference, we can introduce ambiguity in an objective setting. We can then express the patient's decision problem in formal terms: the established drug is represented as the singleton lottery that yields success with probability 0.3 and the experimental surgery as the set of lotteries that yield success with probabilities between 0.2 and 0.45.

Even without understanding the states underlying her choices, the decision maker may still understand how her choices affect consequences, which are the ultimate objects of her utility. She

[1]An important variation is the model of Anscombe and Aumann (1963), where the consequences are a mixture space, which we subsume in our discussion of general subjective Savage theory.

can understand how an option might make her feel, without understanding the causal mechanism or act that delivers that feeling. Without access to a state space, the agent forms some boundaries on the possible consequential probabilities associated with each choice. These restrictions are captured as sets of lotteries. Such restrictions on the space of consequential uncertainty seem especially plausible when summary information is given to the decision maker by an expert, like the doctor in our introduction.

The standard subjective approach does not assume exogenous ambiguity: insofar as ambiguity exists, it is meaningful only in the mind of the agent. This austere view makes no additional assumptions of the world outside the agent's mind. While such parsimony is theoretically elegant, we believe there are compelling reasons to allow objective ambiguity.

First, introducing objective ambiguity as a set of lotteries arms us with more detail, and this detail can capture realistic features of the decision problem. Fully subjective theories provide no device for the agent to incorporate outside information, or lack thereof, about uncertainty into her decision making. For example, the information provided by her doctor comprises an important part of our decision maker's problem.

Second, subjective utility theory has difficulty distinguishing situations without any ambiguity and situations where the decision maker resolves ambiguity by selecting a single probability. For example, the maxmin expected utility model (Gilboa and Schmeidler 1989) represents attitude towards ambiguity by a set of multiple priors. The same set of priors also represents the existence of ambiguity. This confounds the presence of ambiguity with the agent's resolution of that ambiguity. For example, if she obeys the Savage axioms, we cannot distinguish whether ambiguity was resolved in a probabilistically sophisticated manner or there was never any ambiguity to be resolved in the first place. These features are easily separated in a model where the existence of ambiguity is exogenous and independent of its resolution. By starting with assumed sets of lotteries, our model forces this separation by fiat.

Roughly speaking, we propose the following utility representation $U(A)$ for a set A of lotteries:

$$U(A) = \frac{\int_A \phi \circ u \, d\mu}{\mu(A)},$$

where u is a standard affine expected utility function on single lotteries, μ is a probability measure on sets of lotteries, and ϕ is an increasing transformation applied to u. This decision maker considers all of the relevant lotteries in A when making her decision, with their relative consideration fixed across sets by a measure over all lotteries. Her attitude to ambiguity is captured by the transformation ϕ and her weighting μ.

This paper contributes to a recent literature which equips the decision problem with exogenous ambiguity through sets of lotteries or priors. Jaffray (1989) first introduced exogenous ambiguity over lotteries by defining preferences over nonadditive belief functions. Two other papers, independent of and contemporaneous with this one, also use sets of objective lotteries to model ambiguity, but arrive at distinct utility representations. Olszewski (2007) extends the axioms of Dekel (1986), most notably betweenness and dominance, to sets of lotteries and characterizes a generalized form of α-maxmin utility, which evaluates a set by a convex combination of its minimal and maximal elements. Stinchcombe (2003) presents a novel dual formulation of the Expected Utility Theorem on the mixture space of sets. An advantage of these linear approaches is that the resulting representations are well defined over lower dimensional sets, which our measure theoretic approach must finesse. Finally, Ghirardato (2001), Jaffray and Jeleva (2004), Mukerji (1997), and Nehring (1999) consider multi-valued acts which map states of the world to sets of consequences, while Viero (2006) considers acts which map to sets of lotteries.

Another recent strand of research enriches the subjective model with exogenous information about the priors on the state space (Gajdos, Hayashi, Tallon, and Vergnaud 2006, Gajdos, Tallon, and Vergnaud 2004, Giraud 2006, Wang 2003).[2] These models, and an analog of the second order prior model, are experimentally tested and compared by Hayashi and Wada (2007). A possible technical reconciliation between our approach and these is a generalized form of probabilistic sophistication, where an ambiguous act is evaluated by its induced set of distributions over consequences. This reconciliation and its limitations are further discussed when the model is formally introduced in Section 2. These models take the set of lotteries or priors as exogenous; Ghirardato, Maccheroni, and Marinacci (2004), Nehring (2001), and Siniscalchi (2006) suggest various behavioral methods to identify the perceived priors in subjective models.

Defining preference over sets has a rich tradition in the economic literature of choice under ignorance, starting with Arrow and Hurwicz (1972). There, the decision maker chooses between finite sets of sure consequences, but has no further control as to which alternatives are eventually selected from these sets. This interpretation is closely related to ours, since the decision maker knows only that some set of objects or lotteries is possible and has no further information on how nature will select a particular element in the set. On the other hand, Kreps (1979) uses sets of sure consequences as a way of identifying a preference for flexibility. He interprets sets as menus; the decision maker will eventually choose an available option from the menu. A desire for larger menus suggests a desire to keep one's options open.

More recently, Dekel, Lipman, and Rustichini (2001) and Gul and Pesendorfer (2001) use menus of lotteries to model flexibility and commitment. They view sets as menus of stochastic choices, extending the interpretation over determinate sets due to Kreps (1979). We treat sets as information, extending the interpretation due to Arrow and Hurwicz (1972). At a technical level, neither of their representations are special cases of ours, nor does our utility function constitute a special case of theirs. Most notably, no form of independence over non-singleton sets is imposed here.

Our specific representation has formal and interpretive antecedents. A version of a general representation for conditional expectation was already proven and applied to Jeffrey's (1965) syntactic theory of decision by Bolker (1966, 1967), who deserves original credit for the mathematical result.[3] While the mathematical domain of Bolker's result is somewhat different than ours, the substantive

[2]Our approach is particularly related to that of Giraud (2006), who identifies capacities over priors.

[3]We are extremely indebted to Larry Epstein for bringing Bolker's work to our attention, and thank Chris Chambers for subsequent references on the Jeffrey model.

assumptions are essentially similar. The technical differences between Bolker's result and ours will be explained more comprehensively when we present our theorem.

Our interpretation of the mathematical result differs from Jeffrey's theory of joint desirability and probability over logical propositions or sentences, which is not a treatment of ambiguity. While seemingly established in formal philosophy, Jeffrey's framework is less familiar in economics, and it is unclear whether Jeffrey's propositions should be interpreted as consequences, as states, or as acts.[4] Jeffrey's probability on propositions is the only component of his model which involves uncertainty and is not allowed to vary, whereas we consider the space of *all* possible lotteries on consequences. Perhaps our theory can be viewed as a form of Jeffrey's that considers "propositions" regarding risk: both theories have preferences for information, captured as a sentence about the world or as a set of lotteries, and both implement some form of conditional expectation. That said, we suspect Jeffrey might object to our basic model and primitives: "I take it to be the principal virtue of the present theory, that it makes no use of the notion of a gamble or of any other causal notion (Jeffrey 1965, p. 147)."

Our expression of ambiguity aversion through nonlinear aggregation of expected utilities is closely related to the literature which links ambiguity aversion to second order risk aversion. Segal (1987, 1990) first modeled ambiguity using objective two-stage lotteries and pioneered the conceptualization of ambiguity aversion as a failure to reduce objective compound lotteries. More recently, (Ergin and Gul 2004, Klibanoff, Marinacci, and Mukerji 2005, Nau 2006) study subjective models where first and second order uncertainties are two dimensions of a product state space and derive transformations of expected utility functions. The spirit of Segal's approach is also present in work by Halevy and Feltkamp (2005), who model ambiguity as the bundling of correlated risky prospects, and by Giraud (2006), who proposes a nonadditive second order belief. Halevy (2007) reports experimental evidence connecting ambiguity aversion and reduction of compound lotteries and provides a nice categorization and comparative test of the different proposed forms of compounding.

While our representation and its associated interpretation are complementary, our domain is distinct from these studies. We examine *sets* of lotteries, not two-stage *lotteries* over lotteries. Referring back to the Ellsberg urn, for example, the subject's given information is perhaps more transparently modeled as a set of possible distributions of colors ("between zero and sixty yellow balls") than as a compound process of distributions over distributions of colors.

Thus, a corollary contribution of the model is its alternative axiomatic foundation for the second order approach to ambiguity, which includes no explicit or verifiable second order uncertainty of any sort. The decision maker is not asked to rank bets on which lottery actually generates risk.[5] Since they are outside our purview, we have nothing directly to say about reduction of compound lotteries. This is arguably a comparative virtue: the uncertainty over lotteries is produced without appeal to second order measurement devices. Insofar as "second order" uncertainty exists, it is as an artifact of the suggested utility representation, rather than as a primitive assumption of

[4]Nonetheless, there is at least one application of the theory to welfare economics (Broome 1990).
[5]A recent paper by Seo (2006) calibrates the second order prior using lotteries over acts.

the model's domain. This benefit comes at a cost. The use of direct second order measurement identifies the nonlinear aggregator ϕ modulo positive affine transformations. Without this direct measurement, we cannot achieve this level of uniqueness and allow a third degree of freedom.

Aside from the alternative domain and axiomatic scheme, there are two substantive contributions of this model over other approaches to second order uncertainty. First, in the most general representation, preferences over unambiguous options are left completely general. This allows for nonexpected utility as the integrand of the second order belief, and accommodates behavior like the Allais paradox over purely risky decisions.[6] The second is the introduction of a novel concept of luck, defined through the stochastic dominance order on the second order belief. This provides an additional method of comparing ambiguity aversion. Its value becomes especially salient given the first point: without independence over risky choices, ϕ is no longer a meaningful component of the representation. Nonetheless, the comparison of μ is still applicable in this model, allowing us to compare individuals' ambiguity aversion, even if they violate expected utility over risk.

In the next section, we formally introduce the primitives of our theory. Section 3 contains our main representation: the decision maker integrates a transformed expected utility with respect to a second order measure, conditioning on the objective set of lotteries. Section 4 provides a definition of comparative ambiguity aversion and conducts some comparative statics.

2 An objective model of ambiguity

We introduce the domain of preference, a special family of sets of objective lotteries. Preferences over somewhat different families of sets of lotteries are also studied by Olszewski (2007), who provides two representations for all closed sets and for all closed convex polyhedra, and Stinchcombe (2003), who considers closed convex sets.

The finite set X denotes the set of deterministic outcomes. Let ΔX denote the set of lotteries on X, endowed with the topology of weak convergence, which is induced by the Euclidean metric when ΔX is represented as $\Delta X = \{x \in \mathbb{R}_+^{|X|-1} : \sum_{i=1}^{|X|-1} x_i \leq 1\}$. Let $\Delta^2 X = \Delta(\Delta X)$ denote the set of Borel probability measures on ΔX, or second order measures on X. We refer to elements of ΔX as "lotteries," and reserve the term "measures" for elements of $\Delta^2 X$.

A set is regular is it is equal to the closure of its interior. Our domain of choice is the family of nonempty regular and singleton subsets of ΔX, denoted

$$\mathcal{K}^*(\Delta X) = \{A \subseteq \Delta X : \overline{\text{int}(A)} = A \text{ or } |A| = 1\} \setminus \emptyset.$$

Regular sets are not closed under standard set operations: consider the intersection of two regular sets that meet only at their boundaries. Regular sets are closed under the regularized set operations,

[6] Klibanoff, Marinacci, and Mukerji (2005) also present a version of their result for rank dependent preference.

defined as:

$$\begin{aligned} A \cup' B &= \overline{\text{int}(A \cup B)}; \\ A \cap' B &= \overline{\text{int}(A \cap B)}; \\ A \setminus' B &= \overline{\text{int}(A \setminus B)}. \end{aligned}$$

We will slightly abuse notation and drop the primes: all subsequent set operations are regularized. The basic object of study is a binary relation $\succsim$ on $\mathcal{K}^*(\Delta X)$, with $\succ$ and $\sim$ having the standard definitions.

This domain restriction is a significant one. It excludes, for example, all finite subsets with more than one element. The decision maker must either face ambiguity regarding the probabilities of all consequences or face no ambiguity at all; she cannot know precisely the probabilities of some outcomes but not of others.[7] For example, if the decision maker is unsure of the true probability she will survive a surgical procedure, then she must also be unsure of the odds that the sun will rise in the morning. At the point of modeling, such examples seem less disturbing, since presumably outcomes which are determined and over which the decision maker has no control can be excluded from the description of the consequences. Finally, in some convenient parameterizations of ambiguity, such as epsilon contamination where $A = \{(1-\varepsilon)\ell + \varepsilon x : x \in \Delta X\}$, the set of beliefs is either a singleton or regular.

While regular sets are dense in the family of closed subsets, a weakness of our utility representation is that conditional expectation is not defined on null sets. We could include such sets in our domain, but our measure-theoretic representation would have no bite on them.[8] Their ability to put structure on such sets is a comparative strength of the linear approaches of Olszewski (2007) and Stinchcombe (2003).

Any singleton is unambiguous, because the risk is known and precise. We view this simple definition as a strength of the theory, but also acknowledge it is an artifact of the exogenous nature of ambiguity in the model.

Our model, defined directly over sets of lotteries over consequences, can arguably be treated as a reduction of the Savage model, whose primitives are acts $f : S \to X$ mapping states to consequences. Given a probability assessment μ over states, each Savage act f is naturally associated with its induced distribution over consequences: $\nu = \mu \circ f^{-1}$. If the decision maker is probabilistically sophisticated, in the sense of Machina and Schmeidler (1992), these image lotteries completely characterize her preference. Then an act contains more structure than is required for decision making; all payoff relevant information is captured by its distribution. Similarly, sets of consequential lotteries might be viewed as reduced formulations of ambiguity in a Savage setting if

[7] Another way of thinking about the restriction, pointed out to us by Wojciech Olszewski, is that these sets are of the same dimension as ΔX.

[8] Another facile alternative would be to impose the following axioms separately on lower dimensional components of ΔX, for example each face or edge of ΔX. While this would provide structure on comparing two subsets of the same face, it would not restrict preference across faces.

the relevant information is captured by the *set* of possible distributions induced by an act, given a set of probability assessments over states. For example, in the Ellsberg urn, such a reduction implies that the agent treats yellow and black symmetrically and is indifferent between betting on either. This imposed symmetry seems reasonable in many cases. Ellsberg himself reported, "In our examples, actual subjects do tend to be indifferent between betting on [yellow or black]. . . . the reasons, if any, to favor one or the other balanced out subjectively so that the possibilities entered into their final decisions weighted equivalently (Ellsberg 1961, p. 658)."

In other cases, the translation is more tenuous, as illustrated by the following example due to Takashi Hayashi. Consider two urns: the first contains 100 red or green balls, the second contains 100 red, green, or yellow balls. No further information regarding these urns is known. If the decision maker considers only the induced distributions of acts, then betting on a red ball from the first urn induces the same lotteries as betting on a red ball from the second urn, yet she may prefer to bet on the first urn because there are fewer possible colors. This suggests that some important information might be lost in reducing ambiguous acts to their distributions. The association of ambiguous acts in subjective settings to ambiguous sets in our objective model is a delicate one. A relaxation of distributional reduction is proposed by Gajdos, Hayashi, Tallon, and Vergnaud (2006).

Finally, this model sharply delineates ambiguity as a fixed set of lotteries. In reality, the decision maker may not have such crisp boundaries on the possible lotteries. She may instead consider many sets which might represent the actual ambiguity, and holds a belief on the likelihood of these sets, an element of $\Delta(\mathcal{K}(\Delta X))$. Moreover, if the agent can hold ambiguous beliefs about the consequences ΔX, then she may also hold ambiguous beliefs about the ambiguity, captured as $\mathcal{K}(\Delta(\mathcal{K}(\Delta X)))$, and so on. If these levels of higher order ambiguity can be represented with a single expanded space, then the model loses no generality if the consequence space is properly constructed. In another paper, we construct a space of ambiguous beliefs that also provides a universal consequence space for this model (Ahn forthcoming).

3 Representation

One possible resolution of ambiguity is to focus on the worst possible lottery. Let $u : \Delta X \to \mathbb{R}$ be a utility function on single lotteries and consider the following utility function for a set A:

$$U(A) = \min_{a \in A} u(a).$$

This translates the seminal representation of Gilboa and Schmeidler (1989) to our setting.

Maxmin utility has a clean functional form and crisp axiomatic characterizations. Nonetheless, aside from the minimal lottery, the objective form of this representation ignores all the other lotteries included in a set A. Indeeed, Ellsberg anticipated with dissatisfaction: "In almost no cases . . . will the *only* fact worth noting about a prospective action be its 'security level': the 'worst' of the expectations associated with reasonably possible probability distributions. To choose on a 'maxmin' criterion alone would be to ignore entirely those probability judgments for which there is evidence (Ellsberg 1961, p. 662)."

Partly to mitigate this extreme form of ambiguity aversion, α-maxmin utility takes a weighted combination of the worst and best distributions in a set, and is characterized by Ghirardato, Maccheroni, and Marinacci (2004) in a subjective setting and by Olszewski (2007) in an objective setting similar to ours. While α-maxmin utility improves simple maxmin, it inherits some problems. For example, α-maxmin utility still ignores almost all of the information contained in the set of priors or lotteries; preferences are completely characterized by minimal and maximal elements. Take the lotteries over \$0 and \$100, represented on $[0, 1]$ by their probabilities for \$100. Then α-maxmin utility is indifferent between $[0, 0.5] \cup [0.9, 1]$ and $[0, 0.1] \cup [0.5, 1]$, while the latter might be more intuitively appealing.

We suggest an alternative representation. Consider a utility function u on the single lotteries and a probability measure μ on the Borel subsets of ΔX. We propose the following resolution of ambiguity:

$$U(A) = \frac{\int_A u \, d\mu}{\mu(A)}.$$

The agent conditions her utility on a second order measure, given the information that the set A of lotteries obtains. This incorporates every lottery in A, weighted by the measure μ.

A more general representation would be $U(A) = \int u \, d\mu_A$, where μ_A is some probability measure supported on A. Ours is the special case where μ_A is the conditional measure of μ.[9] Methodologically, one immediate benefit of our specialization is its parsimony: the agent's behavior can be summarized with a single measure rather than a multitude, which seems easier to take to applications and to falsify with data. Substantively, the specialization imposes an irrelevance of additional alternatives across beliefs analogous to Bayesian updating. For example, if the lotteries in A carry more weight or are more salient than those in B, this relation should not depend on the possible lotteries outside of A and B. In other words, if A and B are subsets of C and D, then $\mu_C(A) \geq \mu_C(B)$ should imply $\mu_D(A) \geq \mu_D(B)$. Otherwise, the relative importance of A or B will depend on whether they are packaged as parts of C or as parts of D. As we will discuss later, the measure μ serves as a proxy for the agent's assessment of her general "luck" or the psychological salience of different possibilities in decision making. It does not have to be interpreted in terms of statistical likelihood. In fact, the decision maker lacks any objective information about the background likelihoods of different lotteries, so we see no reason these should change from choice to choice. Finally, the general representation seems to preclude any meaningful identification of μ_A without further restrictions. For example, with a fixed utility function u, any ν_A generating the same mean of u will generate the same preferences as μ_A. Letting u vary, the problem becomes even more severe. On the other hand, while there are multiple (u, μ) pairs which can represent the same preferences, for any fixed u, the belief μ, hence the conditional belief μ_A, is identified uniquely under the additional structural assumption. Moreover, even without a fixed u, we can identify both components up to three scalar degrees of freedom.

[9]We thank a referee for suggesting this comparison.

We now introduce an axiomatic characterization of the representation. The first axiom is standard.

Axiom 1 (Weak order)**.** $\succsim$ is complete and transitive.

Continuity over sets is often defined with respect to the Hausdroff metric. The proposed utility violates Hausdorff continuity when defined over all closed sets of lotteries. Suppose the lotteries are over two outcomes, winning \$0 or \$100. This set of lotteries can be represented as $[0, 1]$, indexed by the probability of winning \$100. Set $u(x) = x$ and μ to the Lebesgue measure. Then the sets $A_\delta = [\frac{1}{4} - \delta, \frac{1}{4} + \delta] \cup [\frac{3}{4} - 2\delta, \frac{3}{4} + 2\delta]$ and $B_\delta = [\frac{1}{4} - 2\delta, \frac{1}{4} + 2\delta] \cup [\frac{3}{4} - \delta, \frac{3}{4} + \delta]$ both converge to the doubleton $\{\frac{1}{4}, \frac{3}{4}\}$ as $\delta \to 0$. Yet $U(A_\delta) = \frac{7}{12}$ and $U(B_\delta) = \frac{5}{12}$ for all δ, so continuity fails. The main problem is that $\{\frac{1}{4}, \frac{3}{4}\}$ is a Lebesgue null set, so its conditional expectation is undefined. Our domain of $\mathcal{K}^*$ limits attention to regular sets and singletons and excludes pathological sets null sets. Of course, we include the singletons to retain unambiguous choices.[10]

Putting the representation aside, Hausdorff continuity is arguably inappropriate when sets are interpreted as information. First, Hausdorff distance is an extension of the primitive metric on the singletons. This extension seems sensible when sets are interpreted as menus from which the decision maker will later decide, in which case distances between sets should be inherited from distances between final objects. But, this topology seems more tenuous in our interpretation of sets as information from which nature will decide. Then nature's choice might bear no relationship to the decision maker's preference. Second, once we consider the selection of the chosen lottery as outside the decision maker's control, the uncertainty can depend on cardinalities or sizes which are not preserved under Hausdorff limits. For example, suppose we had restricted attention to finite sets of odd cardinality, which the decision maker evaluates by their median expected utility. As before, the set of winning probabilities is parameterized as $[0, 1]$. Then $A_\delta = \{0, \frac{1}{3}, \frac{2}{3} - \delta, \frac{2}{3}, \frac{2}{3} + \delta\}$ converges in Hausdorff distance to $A = \{0, \frac{1}{3}, \frac{2}{3}\}$ as δ diminishes, yet each $A_\delta \succsim \{\frac{1}{2}\}$ while the limit $A \prec \{\frac{1}{2}\}$. There are many more good possibilities near $\frac{2}{3}$ in A_δ which disappear at the limit. The prior example with intervals was essentially a continuous illustration of the same principle: the relative sizes of the good and worse intervals disappear at the limit. This property means many reasonable notions of the "average" or "expectation" will fail Hausdorff continuity. In fact, we will invoke an alternative measure theoretic notion of continuity to explicitly control for differences in size.

Let λ denote the Lebesgue measure on ΔX. If μ is absolutely continuous with respect to Lebesgue measure, this will be denoted $\mu \ll \lambda$. Symmetric set difference is noted by $A \triangle B = (A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A)$. The next condition asserts that, as the measure of the difference of two sets converges to zero, the sets also converge in preference.

Axiom 2 (Lebesgue continuity)**.** Suppose $\lambda(A) > 0$ and $\lambda(A_n \triangle A) \to 0$. $A_n \succsim B$ for all n implies

[10] Rather than refining the domain of choice, a different approach might invoke a lexicographic probability system (Blume, Brandenburger, and Dekel 1991) of measures. Such an approach would be technically involved, since the number of null sets is very large.

$A \succsim B$ and $B \succsim A_n$ for all n implies $B \succsim A$.[11]

Lebesgue continuity is different from the more standard Hausdorff continuity; there are sequences of sets which converge in Lebesgue measure but do not converge in Hausdorff distance, and vice versa. Given the interpretation of sets as information, Lebesgue continuity, which is based on volume and size rather than on distance, seems more palatable, especially given our domain restriction.

Lebesgue continuity provides no restrictions on the singletons. The following axiom supplements Lebesgue continuity in two ways. First, it guarantees continuity on the purely risky choices. Second, it links the preferences over unambiguous singletons to preferences on nearly unambiguous choices.

Axiom 3 (Downward Hausdorff continuity)**.** Suppose A_n converges to $\{a\}$ in the Hausdorff metric. $A_n \succsim B$ for all n implies $\{a\} \succsim B$ and $A_n \precsim B$ for all n implies $\{a\} \precsim B$.

This axiom is somewhat weaker than the standard Hausdorff continuity condition, which assumes the implication for all Hausdorff convergent sequences of sets. That said, the Hausdorff metric is actually inessential for the axiom's substance. Instead, Axiom 3 can be decomposed into the following two conditions, without reference to Hausdorff distance: first, $\succsim$ is continuous with on the singletons in the weak topology; second, if A_n is a decreasing sequence of sets such that $\bigcap_n A_n = \{x\}$ and $A_n \succsim (\precsim)\, B$, then $\{x\} \succsim (\precsim)\, B$. The reason this axiom is required is because Lebesgue continuity has no bite for unambiguous sets, which are null. The first component provides appropriate continuity for the singletons, while second component links preferences on the regular sets to preference on the singletons through monotone convergence. The Hausdorff metric simply provides a notationally convenient summary of both parts, since the Hausdorff distance on single lotteries coincides with the weak topology and a decreasing sequence of sets intersects at a point $\{x\}$ if and only if its radius, hence its Hausdorff distance from $\{x\}$, becomes arbitrarily small.

Axiom 4 (Disjoint set betweenness)**.** Suppose A, B are regular and disjoint. $A \succsim B$ implies $A \succsim A \cup B \succsim B$ and $A \succ B$ implies $A \succ A \cup B \succ B$.

Gul and Pesendorfer (2001) assume a similar axiom in their work on temptation and self control.[12] Set betweenness carries an obviously distinct interpretation here, since our sets are not menus. In the context of temptation, set betweenness relaxes the following modularity condition: $A \succsim B$ implies $A \sim A \cup B$, which precludes temptation. In the context of ambiguity, set betweenness relaxes exactly the opposite direction, implied by maxmin utility: $A \succsim B$ implies $A \cup B \sim B$.

Given weak order, the axiom may be equivalently written as $A \sim B$ implies $A \sim A \cup B \sim B$ and $A \succ B$ implies $A \succ A \cup B \succ B$ for disjoint regular A, B.[13] In words, the strict part says that the decision maker does not ignore any available information, and every possible lottery carries some weight in her mind. If a strictly preferred and disjoint set A of lotteries is added to the set B,

[11] Lebesgue continuity replaces a divisibility axiom which was assumed in an earlier version of this paper.

[12] Ours is technically weaker in one sense, applying only to disjoint unions; it is stronger in another, preserving both weak and strict preference.

[13] We thank a referee for suggesting this.

then the decision maker must feel strictly better off. A failure implies that the better component of a union of sets is irrelevant to her preference, which in turn suggests that the decision maker is throwing away some relevant information.

The strict component is particularly biting: maxmin and α-maxmin utility both fail the strict part of the axiom.[14] Suppose $\Delta X = [0,1]$, $u(x) = X$, and $\alpha = 1/2$. Then $[\frac{1}{2}, \frac{3}{4}] \succ [0, \frac{1}{4}] \cup [\frac{3}{4}, 1]$ and these sets are disjoint, in the regularized sense. But $[0, \frac{1}{4}] \cup [\frac{1}{2}, 1] \sim [0, \frac{1}{4}] \cup [\frac{3}{4}, 1]$, i.e. the additional good possibility that the lottery might be in $[\frac{1}{2}, \frac{3}{4}]$ has no effect on her decision making.

Axiom 5 (Balancedness)**.** Suppose A, B, C, D are regular with $[A \cup B] \cap [C \cup D] = \emptyset$ and $A \sim B \succ C, D$ (or $A \sim B \prec C, D$). $A \cup C \succsim B \cup C$ implies $A \cup D \succsim B \cup D$.

Balancedness resembles a set theoretic form of independence. Consider the sets $A \cup C$ and $B \cup C$ in the hypothesis. These two unions share C and differ only in that one has A and the other has B. Then, if we replace C with another set D which shares C's preference relation with A, this preference is preserved. The only data that matter in evaluating these types of unions are their set differences, namely A and B. Their intersection, D in the conclusion, does not affect their relative desirability. Similar intuitions, that two objects should be compared by where they are different, are provided to justify independence and monotonicity in other models. Our axiom is weaker because we add the additional restriction that $A \sim B$.

One might wonder why the stronger and seemingly intuitive condition, that $A \succsim B$ implies $A \cup C \succsim B \cup C$, should fail.[15] Consider the following example. A student concerned with her grade point average would rather take a course where only final grades of "A" are assigned to another course where only passing grades are assigned. On the other hand, she might also prefer to take a course which will assign all letter grades from "A" to "F" to a course where only "A"s and "F"s are assigned. The general intuition is that lotteries in B, which are bad when compared only to those lotteries in A, acquire value as a hedge against the even worse lotteries in C. In certain situations, a lot of good news can be better than a bit of great news. The additional hypothesis that $A \sim B$ controls their independent value, and sharpens the comparison to focus exclusively on their hedging value against C.

Maxmin utility is balanced in a somewhat vacuous manner: $A \cup C \sim B \cup C$ whenever $A \sim B \succ C$. Also, the generalized representation where $U(A) = \int_A u \, d\mu_A$ discussed earlier is not balanced unless there is some version of this representation such that $\mu_A = \mu | A$ for some probability μ on all of ΔX.[16]

These five axioms are necessary and sufficient for the proposed utility representation.

[14] Both do satisfy the weak or indifferent part of the axiom. In fact, Olszewski (2007) shows that α-maxmin utility satisfies a strengthened form of weak set betweenness, which he coins "strong generalized betweenness."

[15] We thank a referee for asking this.

[16] To see this, if the collection $\{\mu_A\}$ cannot be represented as conditional measure, there exist some $\mu_C(A)/\mu_C(B) \neq \mu_D(A) \neq \mu_D(B)$. As long as there is some strict preference between singleton lotteries in these sets, this ratio inequality implies a violation of balancedness.

Theorem 1. *A preference relation $\succsim$ on $\mathcal{K}^*$ satisfies Axioms 1–5 if and only if there exist a continuous $u : \Delta X \to \mathbb{R}$ and a probability measure $\mu \ll \lambda$ on ΔX with full support such that*

$$U(A) = \begin{cases} \dfrac{\int_A u\, d\mu}{\mu(A)} & \text{if } A \text{ is regular} \\ u(x) & \text{if } A = \{x\} \end{cases}$$

is a utility representation of $\succsim$.

Moreover, suppose there exists such a utility representation by (u, μ). Then (v, ν) also represent $\succsim$ if and only if

$$\begin{aligned} v(x) &= \frac{au(x) + b}{cu(x) + d}; \\ \nu(B) &= \mu(B)[c\textstyle\int_B u\, d\mu + d]. \end{aligned}$$

for some numbers $a, b, c, d \in \mathbb{R}$ such that $ad - bc > 0$ and $d = 1 - c \int_{\Delta X} u\, d\mu$.

Proof. See Appendix A.

The utility function u on lotteries is not necessarily affine. No form of independence is imposed, accommodating behavior under risk like the Allais paradox. The only restriction over singletons is continuity; to our knowledge, this is the most general form allowed among models with exogenous sets of lotteries or priors.[17] We consider this an important advantage of this measure-theoretic approach. It is also, to our knowledge, the only identification of a second order prior which does not hinge on expected utility over risk. Of course, adding more structure to the preference will provide additional structure on the utility function, which we will do shortly.

As mentioned in the introduction, the technical content of the theorem rediscovers a mathematical result that was proven earlier by and should be credited to Bolker (1966), who provided a functional characterization for quotients of measures on complete nonatomic Boolean algebras. While the motivations, formal hypotheses, and proofs are different, we do not want to claim any significant technical novelty. At the same time, there are some important formal differences in the results. Bolker begins with a nonatomic algebra: this excludes the singletons and would force the existence of ambiguity in our setting. This exclusion might be natural in a propositional Boolean model for logic or syntax, but the inclusion of atoms is essential in our interpretation of sets as ambiguity to formalize unambiguous choices and to get a handle on the the form of the utility function u. To allow for atoms, we require the additional continuity assumptions. We also derive more structure on the functional equation: the Radon–Nikodym derivative, which becomes u in our setting, is continuous and μ is absolutely continuous with full support. While Bolker does not require continuity of u, within our narrower class of continuous utility indices, our continuity conditions become necessary, as well as sufficient, for the representation.

[17] Olszewski (2007) outlines a version of α-maxmin utility where preferences over singletons satisfy Dekel (1986) betweenness.

While a detailed proof is in the appendix, we briefly outline the main ideas here. First, let Λ_A denote the family of regular subsets of A which are indifferent to A. Λ_A is λ-system, being closed under complementation and disjoint unions, both by disjoint set betweenness. Such systems are recently emphasized in the literature on ambiguity (Epstein 1999, Epstein and Zhang 2001, Zhang 1999). The balancedness axiom provides a natural likelihood relation $\succeq_\ell$ on Λ_A: $S \succeq_\ell T$ if and only if $S \cup B \succsim T \cup B$ for some disjoint $B \prec A$. The likelihood relation satisfies sufficient conditions, due to Zhang (1999), for the existence of a quantitative probability measure P_A which represents $\succeq_l$. The measure can be uniquely extended to the entire class of regular subsets indifferent to A. We then use the extended P_A to construct a signed measure ν_A on all of the regular subsets of ΔX such that the sign of $\nu_A(S)$ identifies whether an arbitrary regular set S is preferred to A: $\nu_A(S) \geq 0$ if and only if $S \succsim A$.

Each ν_A provides information about the underlying preference with respect to a fixed set A, serving as a partial representation of the preference. The next step is to connect these signed measures together to construct a complete representation. Take any three regular sets such that $A \succ B \succ C$ and consider the vector-valued measure $\nu_{ABC} = (\nu_A, \nu_B, \nu_C)$ taking values in $\mathbb{R}^3$. Exploiting the constructed representation properties of the measures' signs, an application of Lyapunov Convexity Theorem demonstrates that the image of ν_{ABC} is spanned by two vectors.[18] This implies that ν_C is linearly determined by ν_A and ν_B. Then the family of signed measures $\mathcal{M} = \{\nu_A : A \text{ is regular}\}$ is spanned by two of its elements, ν and μ. Furthermore, the cone generated by $\mathcal{M}$ is convex. This convexity implies that $\mathcal{M}$ is contained in a half space of the vector space of signed measures. Let ν^* be the measure orthogonal to the boundary of this half space. Some simple algebra demonstrates that the fraction $\nu^*(A)/\mu(A)$ is a utility representation over the regular sets. Then the Radon–Nikodym Theorem implies there exists a measurable real-valued function u on ΔX such that $\nu^*(A) = \int_A u \, d\mu$. This demonstrates the utility representation for regular sets; the representation for singletons is the consequence of a technical convergence lemma and downward Hausdorff continuity.

Finally, the nonparametric domain of choice ΔX considers arbitrary distributions on X and makes no restrictive assumptions on the shape of risk. However, the proof only assumes that ΔX is a compact Polish mixture space. The characterization remains valid even if X is not compact or Polish when restricted to a family of parameterized distributions, provided the parameter space is a compact Polish mixture space. This is true even if the consequence space X is neither compact nor Polish. For example, if the decision maker knows that the risk is normal over all levels of wealth, we can define preference over closed subsets of a compact rectangle $M \times V \subset \mathbb{R}^2$, associated with normal distributions of different means M and variances V, while $X = \mathbb{R}$ is not compact.

Adding independence over unambiguous singletons refines the representation. The following condition is standard.

Axiom 6 (Singleton independence)**.** For all $a, b, c \in \Delta X$ and $\alpha \in (0, 1)$, $\{a\} \succsim \{b\}$ if and only if $\{\alpha a + (1 - \alpha)c\} \succsim \{\alpha b + (1 - \alpha)c\}$.

[18] A form of the Lyapunov Convexity Theorem was also invoked to similar effect by Bolker (1966).

A stronger axiom imposes independence over all sets: $A \succsim B$ if and only if $\alpha A + (1-\alpha)C \succsim \alpha B + (1-\alpha)C$, for all $A, B, C \subseteq \Delta X$. This stronger condition is motivated by indifference to the timing of uncertainty. In this interpretation, the mixture $\alpha A + (1-\alpha)C$ is viewed as the lottery where the set A is realized with probability α and C is realized with probability $1-\alpha$. If the decision maker is indifferent to when uncertainty resolves, she should retain her preference of A to B when they are mixed with another set C. This view is invoked when sets are interpreted as menus; our interpretation of sets as information and the resulting representation preclude this interpretation, since the relative importance to decision making of A versus C, when both sets are possible, is fixed by the probability measure μ. Even keeping the interpretation of mixtures as objective randomization over sets, a decision maker facing ambiguity might very well prefer a specific timing of uncertainty, as pointed out by Epstein, Marinacci, and Seo (2007). For example, consider the uncertainty aversion axiom for Anscombe–Aumann acts by Gilboa and Schmeidler (1989): if $f \sim g$, then $\alpha f + (1-\alpha)g \succsim f$. If $\alpha f + (1-\alpha)g \succ f$, then the decision maker prefers to flip a coin between the two uncertain acts than face either on its own. Similarly, facing two ambiguous sets of lotteries, the decision maker might rather face the set of coin flips between elements of the two sets than face either set on its own.

As the following corollary states, singleton independence provides a two step evaluation of a set of lotteries: first each lottery is linearly aggregated by expected utility, then the entire set is nonlinearly aggregated by μ. This is reminiscent of the compounding approach to ambiguity aversion forwarded in Segal (1987, 1990) and of the transformed expected utility aggregators in Ergin and Gul (2004), Klibanoff, Marinacci, and Mukerji (2005), and Nau (2006).

Corollary 2. *A preference relation $\succsim$ on $\mathcal{K}^*$ satisfies Axioms 1–6 if and only if there exist an affine $u : \Delta X \to \mathbb{R}$, a strictly increasing and continuous $\phi : \mathbb{R} \to \mathbb{R}$, and a probability measure $\mu \ll \lambda$ on ΔX with full support such that*

$$U(A) = \begin{cases} \dfrac{\int_A \phi \circ u \, d\mu}{\mu(A)} & \text{if } A \text{ is regular,} \\ \phi(u(x)) & \text{if } A = \{x\} \end{cases}$$

is a utility representation of $\succsim$.

Moreover, suppose there exists such a utility representation by (u, ϕ, μ). Then (v, ψ, ν) also represent $\succsim$ if and only if

$$\begin{aligned} v(x) &= \alpha u(x) + \beta; \\ \psi(z) &= \frac{a\phi(z) + b}{c\phi(z) + d}; \\ \nu(B) &= \mu(B)[c\textstyle\int_B u \, d\mu + d]. \end{aligned}$$

for some numbers $\alpha, \beta, a, b, c, d \in \mathbb{R}$ such that $\alpha, ad - bc > 0$ and $d = 1 - c\int_{\Delta X} u \, d\mu$.

Imposing singleton independence, we retain standard expected utility as a special case on the

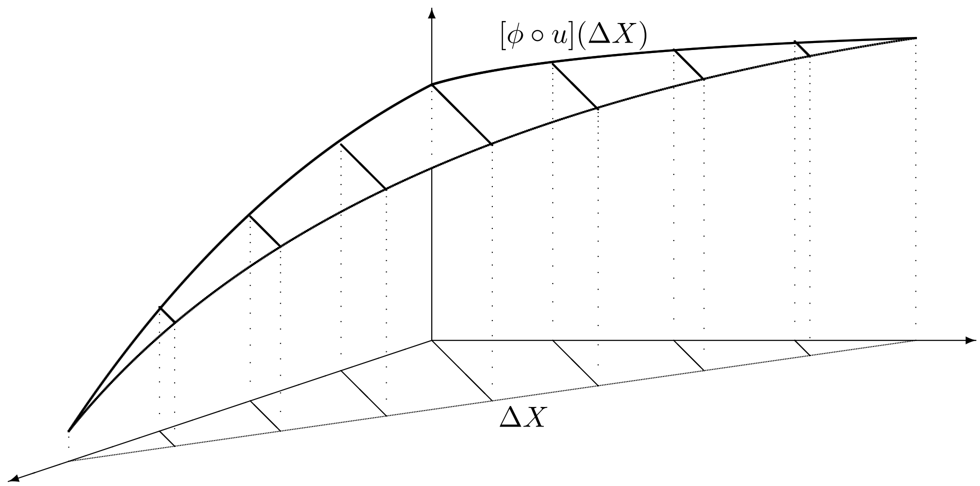

Figure 1: Nonlinear "expected" utility ($\phi \circ u$)

unambiguous singletons. The utility u is a standard affine expected utility function, and ϕ is a transformation that retains the ordinal independence condition on preferences; both u and $\phi \circ u$ produce linear indifference curves on single lotteries. The classic vNM Expected Utility Theorem states there exists *some* affine utility representation, but not that *all* utility representations must be affine. There are nonlinear utility representations of independent preference, as shown in Figure 1.

The classic theory is cardinal with respect to the value function on deterministic outcomes X (up to scale transformations), but is only ordinal with respect to the utility function on lotteries ΔX. A nonlinear monotone transformation is irrelevant in comparing one lottery to another, but becomes very relevant in comparing sets of lotteries. The functional form of utility used to represent preferences over single lotteries is important when we integrate that form over sets of lotteries.

Although related, the curvature of utility in our model is not induced in the traditional sense of direct second order risk aversion on compound or two-stage lotteries. This would suggest a relaxation of the reduction axiom in the space of compound lotteries, about which we have nothing to say. The domain of our model is *sets* of simple lotteries, rather than *lotteries* of simple lotteries. The measure μ over lotteries is fixed by the representation; it is not assumed as a primitive of the theory, nor is it allowed to vary to reflect different second order uncertainties. Instead, the cardinality of $\phi \circ u$ is induced by the agent's preferences over subsets of ΔX.

Thus one part of an agent's attitude towards ambiguity is the transformation ϕ, which captures a cardinal intensity of preference for one type of lottery compared to another. This intensity is important because the utility integrates this intensity with respect to the fixed measure μ. The manifestation of ambiguity aversion as a departure from linearity should be comforting, since it resonates the traditional analysis of risk aversion. We develop this metaphor more carefully later.

The second part of her attitude is determined by her weighting μ. One interpretation of our setting is a game against nature, where nature decides which lottery is actually realized.[19] Then μ is the agent's belief about nature's mixed strategy. If the agent thinks μ puts more probability on worse lotteries, she thinks nature will more likely choose a bad lottery. The measure μ can be interpreted as an agent's assessment of her "luck" in ambiguous situations. The more weight she places on better lotteries, the luckier she believes herself to be. Two agents can share a common utility function u and transformation ϕ, but still have different preferences because one thinks of herself as luckier than the other. The maxmin utility is a particularly salient case where the agent pessimistically believes that nature always chooses the minimal element of a set.[20] This interpretation seem less transparent when ambiguity aversion is solely modeled through ϕ. We recover this transparency by putting the second order belief directly on the objects of utility.

Similar concepts of luck cannot be formalized in a subjective model where the agent takes a fixed weighted average over her priors over states. This is because whether the weighting is optimistic or pessimistic depends on the act being evaluated. For example, if she holds a bet on yellow in the Ellsberg urn, placing more weight on distributions with many yellow balls is optimistic. On the other hand, if she holds a bet on black, the same weighting would be considered pessimistic. Only by putting the second order beliefs directly on consequences can the decision maker be considered optimistic or pessimistic. Finally, we note that this interpretation of μ does not depend on independence; the measure still reflects pessimism, even if risk is not evaluated with expected utility.

Another interpretation of μ is less wedded to the statistical view of the decision problem as a game against nature. We can interpret μ as a measure of salience or how much attention the agent pays to the various lotteries. Worse lotteries loom larger in the minds of those with a distaste for ambiguity. The measure μ then corresponds to the personal attention given to the possible lotteries, normalized so $\mu(\Delta X) = 1$. Then μ captures the psychological, rather than statistical, weight the decision maker attaches to the various lotteries.

The uniqueness of u modulo affine transformations is standard. The uniqueness of the transformation ϕ and the probability measure μ are strictly weaker than that achieved in some other representations. This is due to the different kind of information being elicited. Since we view the direct measurement of second order beliefs as artificial, we do not allow bets on the space ΔX and have no method to directly elicit the decision maker's second order belief. In contrast, the models of Ergin and Gul (2004) and Klibanoff, Marinacci, and Mukerji (2005) identify an exact second order measure μ on first order priors. But this identification benefits from these models' primitives, which ask the decision maker to assess bets on the correct prior. In contrast, we do not access this full range of bets to identify beliefs; the decision maker's probabilistic payoff under lottery x is never supplemented or changed. For example, our decision maker is not asked to compare betting a dollar on the event that the actual lottery generating outcomes is in the set A to betting on the event that the lottery is in another set B. Demanding such hypothetical comparisons

[19] This interpretation is advanced more explicitly by Olszewski (2007).

[20] Of course, this extreme sense of bad luck cannot actually be defined as a probability measure over nature's behavior.

would provide another measurement device that we suspect would obtain uniqueness, but one of our motivations for developing this domain is the artificiality of such contingent comparisons. We feel one of the model's strengths is the lack of such forced comparisons, and in deriving the second order uncertainty purely as an artifact of the utility representation, rather than embedding it into the primitives of the model. This also explains the looser identification. Instead of linear transformations, which have two degrees of freedom, we allow fractional linear transformations with three degrees of freedom, since d is defined as a function of c. Moreover, the second order belief is not identified uniquely, but allowed a degree of freedom, for similar reasons. At a technical level, this is not entirely surprising, the more information we have about the decision maker, such as her preference over independent bets on the lotteries, the more we can identify.

4 Comparative ambiguity aversion

In this section, we develop tools to compare ambiguity aversion across individuals in our objective setting.[21] We begin by providing a behavioral notion of comparative ambiguity attitude.

Definition 1. The relation $\succsim_1$ is *locally more ambiguity averse at* A than $\succsim_2$ if

$$A \succsim_1 \{a\} \Rightarrow A \succsim_2 \{a\}$$

and

$$A \succ_1 \{a\} \Rightarrow A \succ_2 \{a\},$$

for all $a \in \Delta X$.

The relation $\succsim_1$ is *(globally) more ambiguity averse* than $\succsim_2$ if it is locally more ambiguity averse at all $A \in \mathcal{K}^*$.

This definition is analogous to Epstein's (1999) definition of comparative ambiguity aversion. He considers $\succsim_1$ more ambiguity averse than $\succsim_2$ if for every arbitrary act f and any unambiguous act g, $f \succsim_1 g$ implies $f \succsim_2 g$ and $f \succ_1 g$ implies $f \succ_2 g$, where an act is considered unambiguous if it is measurable with respect to a λ-system of unambiguous events. The definition by Ghirardato and Marinacci (2002) is identical, except they further restrict g in the hypothesis to be a constant function. The two definitions disagree on what exactly constitutes an unambiguous act. Here, singleton lotteries directly take the place of λ-measurable or constant acts, allowing us to finesse the issue. With the subjective definitions proposed by Epstein and by Ghirardato and Marinacci, our definition shares the virtue of being applicable across different utility representations for the decision maker.

While the local definition of ambiguity aversion does not reference the cardinal utilities of the decision makers over X, a consequence of the global definition is that if one decision maker is globally more ambiguity averse than another, then both share the same restricted preferences on

[21] A working version of this paper included a notion of objective ambiguity neutrality; details are available from the author upon request.

singleton lotteries. If X represents levels of wealth, if we can compare two agents' global ambiguity attitudes, the agents must have the same cardinal utility for wealth. So the ordering of ambiguity aversion is coarser than the ordering of risk aversion.

Proposition 3. *If $\succsim_1$ is globally more ambiguity averse than $\succsim_2$, then $\succsim_1|_{\Delta X} = \succsim_2|_{\Delta X}$*

Proof. This follows directly from restricting the definition to singletons. □

If two decision makers disagree on the desirability of uncertain prospects, it is because they have different cardinal tastes over the sure consequences or because they have different reactions to the size of ambiguity. Our definition of comparative ambiguity aversion separates the effects of risk and ambiguity on decision making by fiat. Ghirardato and Marinacci consider such separation desirable and carefully delineate conditions where it is implied by their definition in a Savage framework. The immediate separation here is an artifact of our objective domain of lotteries, which accesses a rich linear structure not immediately available in a Savage domain.

Given that two agents share risk preferences, we can consider a_1 and a_2 their respective ambiguity-free equivalents to A if $\{a_1\} \sim_1 A$ and $\{a_2\} \sim_2 A$.[22] Ambiguity-free equivalents are conceptually similar to certainty equivalents in the theory of risk aversion. An agent is more risk averse than another if the other's monetary certainty equivalent for a lottery is greater than her own certainty equivalent. Here, we replace the natural ordering on money with the preference ordering on singletons. Then the first is more ambiguity averse at A than the second if she prefers the second's ambiguity-free equivalent a_2 to her own equivalent a_1: $\{a_2\} \succsim_1 \{a_1\}$. She can be more ambiguity averse for some sets but less ambiguity averse for others, in the same way she can be more risk or less risk averse for different lotteries.

Remember our discussion of the two sides of ambiguity aversion: the transformation ϕ reflecting a cardinal utility towards gambles and the probability assessment μ reflecting an attitude about one's luck. To conduct comparative statics, we isolate each effect by keeping the other fixed. We begin by fixing the measure μ and comparing the curvature of ϕ. In the theory of risk, more curvature corresponds to more risk aversion. Similarly, in our theory more curvature corresponds to more ambiguity aversion. For a fixed risk profile, we let $\succsim_{\text{MMEU}}$ refer to the corresponding maxmin expected utility.

The fraction $-\frac{\phi_i''(x)}{\phi_i'(x)}$ resembles the Arrow–Pratt coefficient of absolute risk aversion. Fixing a probability μ, it similarly serves as a quantitative measure of ambiguity aversion. We mentioned earlier that maxmin utility only barely fails to meet disjoint set-betweenness. In fact, it is a limit case of our representation. As this measure approaches infinity, the agent's preferences approach maxmin utility. Her relative distaste for worse lotteries increases, and she wishes more and more to avoid sets that include such lotteries. This is hardly surprising, given similar results in the theory of risk. Moreover, analogous results in subjective settings are provided by Klibanoff, Marinacci, and Mukerji (2005).

[22]More generally, we can consider the entire set $\{a \in \Delta X : \{a\} \sim A\}$ as the ambiguity-free equivalent. This set is the restriction of a hyperplane to ΔX if $\succsim$ meets singleton independence.

To consider limits in the space of possible preferences, we define a topology on the space of preferences. For a fixed set $A \in \mathcal{K}^*$, let

$$d_A(\succsim, \succsim') = d(\{a : \{a\} \sim A\}, \{a : \{a\} \sim' A\})$$

recalling d is the Hausdorff distance.[23] We now say $\succsim_n \to \succsim$ if $d_A(\succsim_n, \succsim) \to 0$ for all $A \in \mathcal{K}^*$. So, a sequence of preferences converges if the respective ambiguity-free equivalents for A converge for any set $A \in \mathcal{K}^*$.[24]

Proposition 4. *Suppose $\succsim$ and $\succsim'$ have representations (u, ϕ, μ) and (u, ϕ', μ), as in Corollary 2. Then $\succsim$ is more ambiguity averse than $\succsim'$ if and only if $\phi = h \circ \phi'$ for some concave and strictly increasing $h : \mathbb{R} \to \mathbb{R}$.*

Moreover, suppose $\{\succsim_n\}$ have representations $\{(u, \phi_n, \mu)\}$. If each ϕ_n is twice differentiable and, for all $x \in \Delta X$,

$$\min_x -\frac{\phi_n''(x)}{\phi_n'(x)} \to \infty,$$

as $n \to \infty$, then $\succsim_n \to \succsim_{MMEU}$.

Proof. The first part of the proposition follows almost directly from Jensen's Inequality. For the second part, make u positive by adding a sufficiently large constant. The result follows by taking a subsequence $\phi_{n(m)}$ with $\phi_{n(m)}$ a concave transformation of $-e^{-mx}$, then taking $m \to \infty$. □

Now we fix the transformation ϕ and vary the probability μ. Recall μ captures the agent's perception of her luck. The partial order of stochastic dominance formalizes what it means for one agent to consider herself "luckier" than another. A measure stochastically dominates another precisely if it puts more weight on more desirable lotteries. We let $\Delta A = \{\mu \in \Delta X : \mu(A) = 1\}$ for any set $A \subseteq X$. Since u is continuous, the set of minimizers over a compact set A is closed. Let $\Delta(\arg\min_{x \in A} u(x))$ denote the set of probabilities concentrated on the (possibly multiple) minimizers of u. When $\mu(A) > 0$, $\mu|A$ is the conditional probability defined by $[\mu|A](S) = \frac{\mu(A \cap S)}{\mu(A)}$.[25]

Proposition 5. *Suppose $\succsim$ and $\succsim'$ have representations (u, μ) and (u, μ'), as in Theorem 1. If $\mu|A$ stochastically dominates $\mu'|A$ with respect to the lattice $\succsim|_A$, then $\succsim$ is locally more ambiguity averse at A than $\succsim'$.*[26]

Moreover, suppose $\{\succsim_n\}$ have representations $\{(u, \phi, \mu_n)\}$. If, for all measurable $A \subseteq \Delta X$, $\mu_n|A$ converges weakly to $\Delta(\arg\min_{x \in A} u(x))$, as $n \to \infty$, then $\succsim_n \to \succsim_{MMEU}$.

Proof. This follows directly from definitions. □

[23] The arguments in d are compact by continuity of the preferences and d_A is a semimetric on the the space of preferences.

[24] Notice that this convergence need not be uniform across A.

[25] Since μ has full support, this fraction is always well defined for regular sets.

[26] A measure μ *stochastically dominates* ν with respect to the lattice $\succsim$ if $\int f\, d\mu \geq \int f\, d\nu$ for any function f that is monotone with respect to $\succsim$.

This is another way maxmin utility is a limit case of our representation. The more unlucky an agent considers herself, the closer her behavior is to maxmin preference. She focuses more and more of her attention on the worse lotteries, until the worst lottery becomes the sole criterion for comparison. This method of comparison still works even if singleton independence is not assumed.

A Appendix: Proof of Theorem 1

Recall that λ denotes the $(|X|-1)$-dimensional Lebesgue probability measure on ΔX. We begin by proving a useful technical lemma. The lemma makes Corollary 2 an immediate consequence of Theorem 1.

Lemma 6. *There is a version $f = \frac{d\mu}{d\lambda}$ (λ–almost everywhere) of the Radon–Nikodym derivative of μ with respect to λ that is continuous at a if and only if $\frac{\mu(A_n)}{\lambda(A_n)} \to f(a)$ whenever $A_n \to \{a\}$ in the Hausdorff metric topology and $\lambda(A_n) > 0$.*

Proof. We first prove sufficiency by contradiction. Suppose $\frac{\mu(A_n)}{\lambda(A_n)} \to f(a)$ if $A_n \to \{a\}$ and $f(x)$ is discontinuous at a for any version f. Then either $\max\{f(x), f(a)\}$ or $\min\{f(x), f(a)\}$ is discontinuous at a; without loss of generality assume the former. Then there exists some $\varepsilon > 0$ such that

$$D_n = \left\{ x : \|x - a\| < \frac{1}{n} \text{ and } f(a) - f(x) > \varepsilon \right\}$$

is nonempty for all n. Furthermore, $\lambda(D_n) > 0$ for all n, otherwise we can find a version $g = f$ almost everywhere with respect to λ with g continuous at a. Then

$$\frac{\int_{D_n} f \, d\lambda}{\lambda(D_n)} < f(a) - \varepsilon$$

for all D_n while $D_n \to \{a\}$. This contradicts our assumption that $\frac{\mu(A_n)}{\lambda(A_n)} \to f(a)$ whenever $A_n \to \{a\}$.

For necessity, suppose there exists a version f that is continuous at a and take any sequence $A_n \to \{a\}$. Fix $\varepsilon > 0$. There is a corresponding $\delta > 0$ such that $|f(x) - f(a)| < \varepsilon$ whenever $\|x - a\| < \delta$. As $A_n \to \{a\}$, there exists some N such that if $n > N$, then $\|x - a\| < \delta$ for any $x \in A_n$. Then for all $n > N$,

$$\begin{aligned} \left| \frac{\mu(A_n)}{\lambda(A_n)} - f(a) \right| &= \left| \frac{\int_{A_n} f \, d\lambda}{\lambda(A_n)} - f(a) \right| \\ &\leq \frac{\int_{A_n} |f(x) - f(a)| \, d\lambda}{\lambda(A_n)} \\ &< \frac{\int_{A_n} \varepsilon \, d\lambda}{\lambda(A_n)} \\ &= \varepsilon. \end{aligned}$$

□

The routine verification of necessity is omitted. The uniqueness claim follows readily from Bolker (1966). We now provide a proof of sufficiency. We begin by demonstrating the existence of a utility function.

Lemma 7. *If $\succsim$ is downward Hausdorff continuous and satisfies disjoint set betweenness, then for each A, there exists a singleton $\{x\} \sim A$.*

Proof. Fix an arbitrary regular set A. We first prove there exists some x^* such that $\{x^*\} \succsim A$. Let Π_n^i denote the regular partition or grid of A induced by the lattice of points whose dimensions are multiples of 2^{-n}. By repeated applications of disjoint set betweenness, at least one element B_1 of the partition Π_1 satisfies $B_1 \succsim A$. If an element B_n of Π_n satisfies $B_n \succsim A$, then a subset $B_{n+1} \subset B_n$ with $B_{n+1} \in \Pi_{n+1}$ must satisfy $B_{n+1} \succsim B_n$. Thus there exists a sequence of sets B_n such that $B_n \succsim A$ for all n. Since the elements of this sequence are decreasing and have arbitrarily small radius, they converge in the Hausdorff metric to a point $\{x^*\}$. By decreasing Hausdorff continuity, $\{x^*\} \succsim A$. Similarly, there also exists some $\{x_*\} \precsim A$. Decreasing Hausdorff continuity also implies that the restriction of $\succsim$ to the singletons is continuous. Then there must exist some point x with $\{x^*\} \succsim \{x\} \succsim \{x_*\}$ such that $\{x\} \sim A$. □

Therefore, each A has an ambiguity-free equivalent $\{x\} \sim A$. Moreover, by Debreu's Theorem, there exists a utility representation u on the singletons, which can be extended to $\mathcal{K}^*$ through Lemma 7 by setting $V(A) = u(\{x_A\})$, the utility of its ambiguity-free equivalent. We extend $\succsim$ to a larger family of sets. We will consider the family of Borel sets modulo λ.[27] Details of the following constructions can be found in (Halmos 1974, pp. 166–169). We write symmetric set difference as $A \triangle B = (A \setminus B) \cup (B \setminus A)$. Let $\mathcal{Z}'$ refer to the quotient $\mathcal{B}(\Delta X)/\mathcal{N}$, where $\mathcal{N}$ is the family of Borel sets with Lebesgue measure zero. So, if two sets differ only on a set of Lebesgue measure zero, $\lambda(A \triangle B) = 0$, then they are considered part of the same equivalence class in $\mathcal{Z}'$. We will remove the equivalence class of the empty set, to make $\mathcal{Z} = \mathcal{Z}' \setminus [\emptyset]$. At times, we abuse notation and refer to the equivalence class $[A]$ by a representative element A; this should not cause any confusion. Let $\pi(A, B) = \lambda(A \triangle B)$, providing a hereditarily separable metric on $\mathcal{Z}$ (Halmos 1974, Theorem B, p. 168).

Lemma 8. *Let A be a Borel subset of ΔX. If $\varepsilon > 0$, there exists a regular set K such that $\lambda(K \triangle A) < \varepsilon$.*

Proof. Fix a Borel set A and $\varepsilon > 0$. Since λ is an outer regular measure, there exists an open set $O \supseteq A$ such that $\lambda(O \setminus A) < \varepsilon/2$. Let $\mathcal{A}$ denote the family of all closed cubes contained in O. This family is a Vitali covering of A.[28] By the Vitali Covering Theorem (Dunford and Schwartz 1957, Theorem 3, p. 212), there exists a sequence of disjoint sets $A_1, A_2, \dots$ such that $\lambda(A \setminus \bigcup_{i=1}^{\infty} A_i) = 0$. Then there exists a finite index n such that $\lambda(A \setminus \bigcup_{i=1}^{n} A_i) < \varepsilon/2$. Let $K = \bigcup_{i=1}^{n} A_i$. K is a

[27] The particular use of Lebesgue measure here is inessential; any nonatomic measure with full support will suffice.

[28] In $\mathbb{R}^{|X|}$, a family $\mathcal{A}$ of closed sets is a *Vitali covering* of A if each set has strictly positive Lebesgue measure and every point in A is contained in sets of $\mathcal{A}$ with arbitrarily small diameter.

finite union of regular sets, hence regular. Since $K \subseteq O$, $\lambda(K \setminus A) \leq \lambda(O \setminus A) < \varepsilon/2$. Thus $\lambda(A \triangle K) = \lambda(K \setminus A) + \lambda(A \setminus K) < \varepsilon$. $\square$

Lemma 9. *There exists a π-continuous extension of $\succsim$ to $\mathcal{Z}$.*

Proof. Fix a nonnull Borel set B. By Lemma 8, there exists a sequence A_n converging to B in Lebesgue measure. Let $V(B) = \lim V(A_n)$. This sequence is convergent as u is continuous and the Lebesgue measure metric is complete. The Lebesgue continuity axiom makes the selection of any particular sequence inessential. Lebesgue continuity also implies π-continuity of this extension. $\square$

Adding the singletons to $\mathcal{Z}$, when necessary, is done in the natural fashion. At times, we will move between the spaces with and without the singletons appended, but this should not cause any confusion.

We will prove that the axioms on the extended preference are sufficient for the utility representation on all of $\mathcal{Z}$. Then the utility will also represent the original $\succsim$ on the restricted domain $\mathcal{K}^*$.

To prove sufficiency, we reference the theory of probability representation on λ-systems of subsets. This theory was developed conceptually in the mathematical foundations of quantum mechanics by Birkhoff and von Neumann (1936) and von Neumann (1955). More recently, such structure is exploited by Zhang (1999) and Epstein and Zhang (2001) to define the algebraic properties of unambiguous events in the Savage state space.

Definition 2. A family Λ of subsets of X is a λ-*system* if:

1. $X \in \Lambda$,
2. $S \in \Lambda$ implies $S^{\complement} \in \Lambda$, and
3. if $A_1, A_2, \ldots \in \Lambda$ are pairwise disjoint, then $\bigcup_{n=1}^{\infty} A_n \in \Lambda$.

The family of indifferent subsets of A, notated as Λ_A, is conveniently a λ-system.

Lemma 10. *The family $\Lambda_A = \{S \subseteq A : S \sim A\}$ is a λ-system (relative to A).*

Proof. The first condition follows from completeness of $\succsim$. If $S \in \Lambda_A$, then its (relative) complement $A \setminus S \sim A$ to satisfy disjoint set betweenness. Closure under finite disjoint unions follows directly from disjoint set betweenness. For any countable disjoint sequence $\{A_n\}_{n=1}^{\infty}$, $\lambda(\bigcup_{n=1}^{N} A_n)$ converges to $\lambda(\bigcup_{n=1}^{\infty} A_n)$ as N goes to infinity because $\lambda(\bigcup_{n=1}^{\infty} A_n)$ is finite. Then the third condition holds by applying disjoint set betweenness to the finite unions $\bigcup_{n=1}^{N} A_n$, then passing the limit to the preferences using Lebesgue continuity. $\square$

There are various results that find sufficient conditions on a qualitative likelihood ranking $\succeq_\ell$ on a λ-system for the existence of a consistent probability measure, for example (Suppes 1966, Theorem 3) or (Krantz, Luce, Suppes, and Tversky 1971, p. 215). We state a recent version by Zhang (1999), which we find to be the most intuitive and transparent. The symbols $\succ_\ell$ and $\sim_\ell$ carry their natural meanings.

Theorem 11 (Zhang)**.** *There exists a unique finitely additive, convex-ranged*[29] *probability measure P on Λ such that $A \succeq_\ell B \Leftrightarrow P(A) \geq P(B)$ for all $A, B \in \Lambda$ if and only if $\succeq_\ell$ satisfies:*

1. *$A \succeq_\ell \emptyset$ for any $A \in \Lambda$.*

2. *$X \succ_\ell \emptyset$.*

3. *$\succeq_\ell$ is a weak order.*

4. *If $A, B, C \in \Lambda$ and $A \cap C = B \cap C = \emptyset$, then $A \succ_\ell B$ if and only if $A \cup C \succ_\ell B \cup C$.*

5. *For any two uniform partitions $\{A_i\}_{i=1}^n$ and $\{B_i\}_{i=1}^n$ of S in Λ, $\bigcup_{i \in I} A_i \sim_\ell \bigcup_{i \in J} B_j$ if $|I| = |J|$.*[30]

6. (a) *If $A \in \Lambda$ and $A \succ_\ell \emptyset$, there is a finite partition $\{A_1, \ldots, A_n\}$ of X in Λ such that:*
      i. *$A_i \subseteq A$ or $A_i \subseteq A^{\complement}$ for all A_i, and*
      ii. *$A \succ_\ell A_i$ for all A_i.*

   (b) *If $A, B, C \succ_\ell \emptyset$, $A \cup C = \emptyset$, and $B \succ_\ell A$, then there is a finite partition $\{C_1, \ldots, C_n\}$ of C in Λ such that $B \succ_\ell A \cup C_i$ for all C_i.*

7. *If $\{A_n\}$ is a decreasing sequence in Λ and $A^* \succ_\ell \bigcap_n A_n \succ_\ell A_*$ for some A^* and A_* in Λ, then there exists N such that $A^* \succ_\ell A_n \succ_\ell A_*$ for all $n \geq N$.*

We can apply this theorem to Λ_A and the resulting quantitative probability has some nice properties in terms of representing $\succsim$ on a restricted domain. We begin by proving a technical step.

Lemma 12. *For all $A \subseteq X$, there exists $A_0, A_1 \subset A$ such that $A_0 \cap A_1 = \emptyset$, $A_0 \cup A_1 = A$, and $A_0 \cup B \sim A_1 \cup B$ for any $B \nsim A$.*

Proof. Let $\overline{A} = \{x \in A : \{x\} \succsim A\}$ and $\underline{A} = \{x \in A : A \succ \{x\}\}$. Then, by an argument similar to the proof of Lemma 7, $\overline{A} \succ \underline{A}$. Fix some $\overline{x} \in \overline{A}$ and $\underline{x} \in \underline{A}$. Let $\overline{A}_\alpha = \alpha\{\overline{x}\} + (1-\alpha)\overline{A}$ and $\underline{A}_\alpha = \alpha\{\underline{x}\} + (1-\alpha)\underline{A}$. Since $\underline{A} \setminus \underline{A}_\beta \prec A$, disjoint set betweenness forces $\overline{A} \cup \underline{A}_\beta \succ A$. Fix $\beta \in [0,1]$. By construction, $A \succ \underline{A}$. Then Lebesgue continuity implies there exists some $\alpha \in [0,1]$ such that $\overline{A}_{\alpha(\beta)} \cup \underline{A}_\beta \sim A$. Moreover, this α is unique by disjoint set betweenness. Let the function $\alpha(\beta) : [0,1] \to [0,1]$ denote this assignment for each β, which is continuous by Lebesgue continuity. Thus $\overline{A}_{\alpha(\beta)} \cup \underline{A}_\beta \sim A$. By disjoint set betweenness, $A \setminus [\overline{A}_{\alpha(\beta)} \cup \underline{A}_\beta] \sim A$ as well.

Recall $f(A)$ is a Lebesgue continuous utility representation of $\succsim$. Fix $B \prec A$; the argument for $B \succ A$ is parallel. Let $F(\beta) = f(\overline{A}_{\alpha(\beta)} \cup \underline{A}_\beta \cup B)$ and $G(\beta) = f(A \setminus [\overline{A}_{\alpha(\beta))} \cup \underline{A}_\beta] \cup B)$. By construction, $G(0) = f(A \cup B)$, $G(1) = f(B)$, $F(0) = f(B)$, and $F(1) = f(A \cup B)$. Letting $H(\beta) = G(\beta) - F(\beta)$, we have $H(0) = f(A \cup B) - f(B) > 0 > f(B) - f(A \cup B) = H(1)$. As H is continuous in α, the Intermediate Value Theorem implies there exists β^* such that $H(\alpha^*) = 0$, i.e.

[29] A set function P on Λ is *convex-ranged* if for all $A \in \Lambda$ and $0 < \alpha < 1$, there exists some $B \subset A$ such that $P(B) = \alpha P(A)$. Notice this is much stronger than asserting that $P(\Lambda)$ is convex.

[30] A partition $\{A_i\}_{i=1}^n$ is *uniform* if $A_i \sim_\ell A_j$ for all i, j.

$\overline{A}_{\alpha(\beta^*)} \cup \underline{A}_{\beta^*} \cup B \sim (A \setminus [\overline{A}_{\alpha(\beta^*)} \cup \underline{A}_{\beta^*}]) \cup B$. By balancedness, showing this indifference relation for a particular B proves it for all such B. □

Lemma 13. *Take any $A \succ B$ with $A \cap B = \emptyset$. There exist finitely additive, convex-ranged, tight*[31] *probability measures P_A and P_B on Λ_A and Λ_B such that:*

1. *$u(S \cup B) = P_A(S)$ is a utility representation of $\succsim$ on $\{S \cup B : S \in \Lambda_A\}$; and*

2. *$u(A \cup T) = -P_B(T)$ is a utility representation of $\succsim$ on $\{A \cup T : T \in \Lambda_B\}$.*

Furthermore, P_A is robust to choice of $B \prec A$ and P_B is robust to choice of $A \succ B$.

Proof. Define the likelihood ordering $\succeq_\ell$ on Λ_A by $S_1 \succeq_\ell S_2$ if and only if $S_1 \cup B \succsim S_2 \cup B$ and similarly for Λ_B.

Step 1: Conditions 1, 2, 3, and 4. Conditions 1 and 2 are immediate consequences of disjoint set betweenness. Condition 3 follows since $\succsim$ is a weak order. Condition 4 follows immediately from balancedness.

Step 2: If A_0 is a strict subset of A_1, then $A_1 \succ_\ell A_0$. Since Λ_A is closed under disjoint unions, $A_1 \setminus A_0 \sim A$. We have $A \sim A_0 \succ B$, so disjoint set betweenness implies $A \succ A_0 \cup B$. Also by disjoint set betweenness, $A_1 \setminus A_0 \succ A_0 \cup B$ implies $A_1 \cup B = (A_1 \setminus A_0) \cup A_0 \cup B \succ A_0 \cup B$. By definition, this means $A_1 \succ_\ell A_0$.

Step 3: Suppose $A_1 \cap A_2 = A_1' \cap A_2' = \emptyset$, $A_1 \sim A_2$, and $A_1' \sim A_2'$. If $A_1 \succ_\ell A_1'$, then $A_1 \cup A_2 \succ_\ell A_1' \cup A_2'$. By Lemma 12, there exists disjoint B_1, B_2 such that $B_1 \cup B_2 = B$ and $A_1 \cup B_1 \sim A_1 \cup B_2$. By balancedness on B_1, B_2, this implies $A_2 \cup B_1 \sim A_2 \cup B_2$. The definition of $\succeq_\ell$ implies $A_1 \cup B \sim A_2 \cup B$. Then balancedness on A_1, A_2 implies $A_1 \cup B_1 \sim A_2 \cup B_1$. Transitivity of $\succsim$ on the previous indifference relations implies $A_1 \cup B_1 \sim A_2 \cup B_2$. Then disjoint set betweenness forces $A_1 \cup B_1 \cup A_2 \cup B_2 = A_1 \cup A_2 \cup B \sim A_1 \cup B_1$. A parallel argument establishes that $A_1' \cup A_2' \cup B \sim A_1' \cup B_1$. Since $A_1 \succ_\ell A_1'$, we have $A_1 \cup B \succ A_1' \succ B$. Then balancedness implies $A_1 \cup B_1 \succ A_1' \cup B_1$. Thus $A_1 \cup A_2 \cup B \succ A_1' \cup A_2' \cup B$, i.e. $A_1 \cup A_2 \succ_\ell A_1' \cup A_2'$.

Step 4: For any n, there exists a uniform partition $\{B_i\}_{i=1}^{2^n}$ of B. The proof is by induction. By divisibility, there exists a partition $\{B_1, B_2\}$ of cardinality 2. Now, suppose there exists a uniform partition $\{B_1, \ldots, B_{2^n}\}$. By applying divisibility to each B_k, we can produce a two-element partition $\{B_k^1, B_k^2\}$ of B_k such that $B_k^1 \cup A \sim B_k^2 \cup A$, i.e. $B_k^1 \sim_\ell B_k^2$. Step 3 implies $B_k^1 \sim_\ell B_m^1$ for all k, m; $B_k^1 \succ_\ell B_m^1$ would imply $B_k \succ_\ell B_m$, which would contradict $B_k \sim_\ell B_m$. Thus, $\{B_1^1, B_1^2, \ldots, B_{2^n}^1, B_{2^n}^2\}$ is a uniform partition of cardinality 2^{n+1}.

Step 5: Condition 5. Suppose $\{A_i\}_{i=1}^n$ and $\{A_i'\}_{j=1}^n$ are uniform partitions of A. By the previous step, there exists a sequence of disjoint subsets $B_1, \ldots, B_n$ of B such that $B_i \sim B$ and $B_i \sim_\ell B_j$ for all i, j. Let $B_0 = \bigcup_{i=1}^n B_i$.

We first prove that $A_i \sim A_j'$ for all i, j. To the contrary, assume without loss of generality that $A_1 \succ_\ell A_1'$. Then $A_i \sim_\ell A_1 \succ_\ell A_1' \sim_\ell A_j'$. By two applications of balancedness, $A_1 \cup B_1 \sim A_i \cup B_1 \sim A_i \cup B_i$ and similarly $A_1' \cup B_1 \sim A_j' \cup B_j$. By construction, $A \cup B_0 = \bigcup_{i=1}^n A_i \cup B_i = \bigcup_{j=1}^n A_j' \cup B_j$. But

[31] A set function P is *tight* if for every set $A \in \Lambda$, $P(A) = \sup\{P(K) : K \in \Lambda, K \text{ compact}, K \subseteq A\}$.

disjoint set betweenness iteratively applied to $A_i \cup B_i \succ A'_i \cup B_i$ implies $\bigcup_{i=1}^n A_i \cup B_i \succ \bigcup_{j=1}^n A'_j \cup B_j$, a contradiction.

Now consider $\bigcup_{i=1}^m A_i$ and $\bigcup_{j=1}^m A'_j$ for some $m \leq n$. We have just shown that $A_i \sim_\ell A'_i$ for all i. Then disjoint set betweenness iteratively applied to $A_i \cup B_i \sim A'_i \cup B_i$ implies $\bigcup_{i=1}^m (A_i \cup B_i) \sim \bigcup_{j=1}^m \left(A'_j \cup B_j\right)$, i.e. $(\bigcup_{i=1}^m A_i) \cup (\bigcup_{i=1}^m) \sim (\bigcup_{j=1}^m A'_j) \cup (\bigcup_{j=1}^m B_j)$. Then balancedness implies $(\bigcup_{i=1}^m A_i) \cup B \sim (\bigcup_{j=1}^m A'_i) \cup B$. By definition, this means $\bigcup_{i=1}^m A_i \sim_\ell \bigcup_{j=1}^m A'_j$.

Step 6: Condition 6. We first prove part (a). Fix $A^0 \subseteq \Lambda_A$ such that $A_0 \neq \emptyset$ and let $A^1 = A \backslash A^0$. It suffices to show that there exists a finite partition $\{A_1, \ldots, A_n\}$ of A such that $A_i \subseteq A^0$ or $A_i \subseteq A^1$ and that $A_i \succ_\ell A^0, A^1$. If $A_1 \sim_\ell A_0$, we can use divisibility to split both sets and we are done. So, without loss of generality, suppose $A^1 \succ_\ell A^0$. By Step 4, for any m, we can find a uniform partition $\mathcal{A}^{(m)} = \{A_1, \ldots, A_{2m}\}$ of A^1. By additivity of measure, there must exist $A_i^1 \in \mathcal{A}^{(m)}$ with $\lambda(A_i) < 2^{-m}$. Then there exists a sequence of uniform partitions $\mathcal{A}^{(1)}, \mathcal{A}^{(2)} \ldots$ of A^1 such that $\lambda(A_1^{(m)}) \to 0$. Suppose $A_1^{(m)} \succeq_\ell A^0$ for all m. Then, by Condition 4, $A_1^{(m)} \cup (A^1 \setminus A_1^{(m)}) \succeq_\ell A^0 \cup (A \setminus A_1^{(m)})$. But the left side of the relation is exactly A^1 and the right hand side converges in λ to A. Then Lebesgue continuity implies $A^1 \succeq_\ell A$, a contradiction of Step 2. So, there exists a uniform partition $\{A_1^1, \ldots, A_m^1\}$ such that $A^0 \succ A_i^1$ for all i. The same argument, applied to $A_1^{(m)}$ and A^0, provides a uniform partition $\{A_1^0, \ldots, A_n^0\}$ of A^0 such that $A_1^{(m)} \succ_\ell A_1^0$, a fortiori that $A^1 \succ_\ell A_1^0$. Then the union of the two partitions is the required partition of ΔX.

Part (b) similarly follows from Lebesgue continuity and Step 4.

Step 7: Condition 7. This following immediately from Lebesgue continuity. □

Later in the proof, we require a weighting on all sets indifferent to A. This step is not immediate because the family of all sets indifferent to A does not contain a superset to use as X. We now extend P_A to the entire level set.

The first application of Lemma 13 is in proving an intermediate preference result. In words, if A is preferred to B, we can "calibrate" preference by taking subsets of either the better set A or the worse set B.

Lemma 14. *Suppose $A \succsim B$ and $A \cap B = \emptyset$. If $A \succsim C \succsim A \cup B$, then there exists $B' \in \Lambda_B$ such that $A \cup B' \sim C$. If $A \cup B \succsim C \succsim B$, then there exists $A' \in \Lambda_A$ such that $A' \cup B \sim C$.*

Proof. If $A \sim B$, the result follows immediately from disjoint set-betweenness using A, B as A', B'. So take $A \succ B$ and assume the first case, $A \succ C \succ A \cup B$. Let P_B refer to the measure provided by Lemma 13 and the corresponding representation u. Take any continuous utility representation f of $\succsim$ over all $\mathcal{Z}$, normalized so $f = u$ on $\{A \cup T : T \in \Lambda_B\}$. Then $f(A) = 0$; $f(A \cup B) = -1$; and $-1 \leq f(C) \leq 0$. Recall that Lemma 13 also states that $u(A \cup T) = -P_B(T)$ for any $T \in \Lambda_B$. Since P_B is convex-valued, there exists some $B' \in \Lambda_B$ such that $P_B(B') = -f(C)$. Then $f(A \cup B') = u(A \cup B') = -P_B = f(C)$, which proves the first statement of the Lemma. The proof of the second statement is symmetric. □

Lemma 15. *There exists a finitely additive set function that extends P_A to the level set of A, which is unique up to a scale transformation.*

Proof. Fix $A \in \mathcal{Z}$. If $A \sim \Delta X$, then the extension is already provided by $P_{\Delta X}$. So, we may assume that $A \nsim \Delta X$. We assume $A \succ \Delta X$, the case $A \prec \Delta X$ is entirely analogous. Suppose $B \sim A$. Case 1: There exists $b \in B$ such that $\{b\} \sim B$. Then, since $A \succ \Delta X$, by downward Hausdorff continuity we can assume that there exists a set $B' \subseteq B$ containing b such that $\lambda(B') > 0$. Moreover, by π-continuity, since either $b \in A$ or $b \notin A$, we can assume without loss of generality that $B' \subseteq A$ or $B' \cap A = \emptyset$. If it is the former, set $P_A(B) = \frac{P_A(B')}{P_B(B')}$. If it is the latter, then $B' \cup A \sim A$ by disjoint set betweenness. Then set $P_A(B) = \frac{P_B(B')}{P_{A \cup B'}(B') P_{A \cup B'}(A)}$. Case 2: There exists some $b \notin B$ such that $\{b\} \sim B$. By downward Hausdorff continuity, we can assume there exists a set B' disjoint from B and containing b such that $B' \sim B$. Then find $P_A(B' \cup B)$ by Case 1. Set $P_A(B) = P_{B' \cup B}(B) P_A(B' \cup B)$. To observe finite additivity, observe that if B_1 and B_2 are disjoint and $B_1 \sim B_2 \sim A$, then $B_1 \cup B_2 \sim A$ by disjoint set betweenness and additivity of P_A then follows from the additivity of $P_{B_1 \cup B_2}$. □

In Lemma 13, we constructed a measure that represents preference between certain unions of a fixed set S and the members of Λ_A. We now construct a signed measure whose sign indicates preference relative to a fixed set A.

Lemma 16. *If $A \in \mathcal{Z}$, then there exists a nonatomic, finitely additive, tight signed measure ν_A on $\mathcal{Z}$ such that: $\nu_A(S) \geq 0$ if and only if $S \succsim A$.*

Proof. For a fixed A, let $\bar{L} = \{l \in \Delta X : l \prec A\}$ and $\mathcal{L} = \{L \in \mathcal{Z} : L \subseteq \bar{L}\}$. Similarly, let $\bar{U} = \{l \in \Delta X : l \succ A\}$ and $\mathcal{U} = \{U \in \mathcal{Z} : U \subseteq \bar{U}\}$. Then $\mathcal{L}$ and $\mathcal{U}$ are respectively σ–algebras of $\bar{L}$ and $\bar{U}$, after appending the empty set. Either $A \succsim \Delta X$ or $A \precsim \Delta X$. We will assume the former; the argument for the second case is similar.

Fix any $U \in \mathcal{U}$. By our intermediate calibration result, Lemma 14, there exists some $L \in \Lambda_{\bar{L}}$ such that $U \cup L \sim A$. Set $\nu(U) = P_{\bar{L}}(L)$, where $P_{\bar{L}}$ is produced by Lemma 13. By the construction of $P_{\bar{L}}$, $\nu(U)$ is robust to our choice of L, hence uniquely defined.

Select $U_1, U_2 \in \mathcal{Z}$ with $U_1 \cap U_2 = \emptyset$. There exists some $L_{12} \in \Lambda_{\bar{L}}$ such that $L_{12} \cup U_1 \cup U_2 \sim A$, by Lemma 14. Without loss of generality, assume $U_1 \succsim U_2$. By disjoint set betweenness, $L_{12} \cup U_1 \succsim A$. Then we can apply Lemma 14 again to L_{12} to find $L_1 \subset L_{12}$ with $L_1 \cup U_2 \sim A$. Set $L_2 = L_{12} \setminus L_1$. Another application of disjoint set betweenness forces $U_2 \cup L_2 \sim A$, since $L_1 \cup L_2 \cup U_1 \cup U_2 = (L_1 \cup U_1) \cup (L_2 \cup U_2) \sim A$ and $L_1 \cup U_1 \sim A$. Therefore, ν inherits disjoint additivity from $P_{\bar{L}}$. We now have a finitely additive signed measure ν on $\mathcal{U}$.

Now take any $L \in \mathcal{L}$. If there exists some $U \in \mathcal{U}$ such that $L \cup U \sim A$, let $\nu(L) = -\nu(U)$. On the other hand, if there is no $U \in \mathcal{U}$ with $L \cup U \sim A$, we can use Lemma 14 to produce a subset $L' \in \Lambda_L$ such that there exists $U' \in \mathcal{U}$ with $L' \cup U' \sim A$. Let P_L refer to the measure on Λ_L produced by Lemma 13. Set $\nu(L) = \frac{\nu(U')}{P_L(L')}$. The measure ν inherits disjoint additivity on $\mathcal{L}$ from $\mathcal{U}$ by its construction. This extends ν to $\mathcal{L}$.

We move to any arbitrary $S \in \mathcal{Z}$. If $s \sim A$ for all $s \in S$, set $\nu(S) = 0$. Otherwise, we can express this set as $S = L \cup U$ for some $L \in \mathcal{L}$ and $U \in \mathcal{U}$. Set $\nu(A) = \nu(L) + \nu(U)$. Additivity is immediately inherited from $\mathcal{L}$ and $\mathcal{U}$. We have now extended ν to all of $\mathcal{Z}$.

Verifying the representation claim, suppose $S \succsim A$. Then $(S \cap \bar{L}) \cup (S \cap \bar{U}) \succsim A$. By Lemma 14, we can find a subset $U' \in \Lambda_{S \cap \bar{U}}$ with $(S \cap \bar{L}) \cup U' \sim A$. Since $U' \subseteq S \cap \bar{U}$, $\nu(U') \leq \nu(S \cap \bar{U})$. Recalling the construction, $\nu(S) = \nu(S \cap \bar{L}) + \nu(S \cap \bar{U}) = \nu(S \cap \bar{U}) - \nu(U') \geq 0$. Similar arguments establish that $\nu(S) > 0$ only if $S \succ A$.

The measure ν_A is convex-ranged, hence nonatomic. Tightness is inherited from P_A by construction. □

Up to this point, we have only considered finitely additive measures, sometimes called charges. We now show that they are also countably additive, which is a side consequence of Lebesgue continuity.

Lemma 17. *The measure ν_A is countably additive.*

Proof. Each ν_A is finite and ΔX is a Hausdorff space. Since ν_A is tight by Lebesgue continuity, (Aliprantis and Border 1999, Theorem 10.4) implies ν_A is countably additive. □

Now consider $ca(\mathcal{Z})$, the set of all countably additive finite (signed) measures on $\mathcal{Z}$, which is a topological vector space under the topology of weak convergence. The representation features of ν_A are robust to positive scalar transformations, i.e. $\alpha\nu_A$ has the same properties whenever $\alpha > 0$. The measures constructed in Lemma 16 live in $ca(\mathcal{Z})$. The next result shows that these measures can be spanned by two elements of $ca(\mathcal{Z})$. This base will become the critical part of the representation.

Lemma 18. *The family $\{\nu_A : A \in \mathcal{Z}\}$ is spanned by two measures ν, μ, with μ a probability measure.*

Proof. Take any A, B, C which are not indifferent to each other. We lose no generality by ordering them $A \succ B \succ C$. All the measures ν_A are nonatomic. We can invoke the Lyapunov Convexity Theorem: the range of the vector-valued measure $[\nu_A, \nu_B, \nu_C]$,

$$[\nu_A, \nu_B, \nu_C](\mathcal{Z}) = \{(\nu_A(S), \nu_B(S), \nu_C(S)) \in \mathbb{R}^3 : S \in \mathcal{Z}\},$$

is convex.[32] Take any S with $\nu_A(S) = 0$. By construction, $\nu_B(S) > 0$ and $\nu_C(S) > 0$. By using Lemma 14, we can find $S^* \in \Lambda_S$ with $S^* \cup L_1 \sim B$ and $S^* \cup L_2 \sim C$ for some L_1, L_2 disjoint from S^*. Recalling the representation condition in Lemma 16, $\nu_B(S^*) + \nu_B(L_1) = \nu_B(S^* \cup L_1) = 0$; similarly $\nu_C(S^*) + \nu_C(L_1) = 0$. Now take any other T with $\nu_A(T) = 0$, and assume without loss of generality that $T \cap S^* = 0$. Then, suppose $\nu_B(T) \geq \nu_B(S^*)$. Then $\nu_C(T) \geq \nu_C(S^*)$, by disjoint set betweenness and the representation condition applied again. The same arguments hold for the strict inequality as well. Therefore ν_B and ν_C induce the same ordering on $\{S : S \sim A\}$. Applying Lemma 15 and the uniqueness claim of Theorem 11, this ordering completely determines ν_B and ν_C on this restricted domain up to a scale transformation, i.e. $\nu_B(S) = c\nu_C(S)$ for a positive constant c, across any S with $\nu_A(S) = 0$. Then $\{x \in \mathbb{R}^3 : x_1 = 0\} \cap [\nu_A, \nu_B, \nu_C](\mathcal{Z})$ is contained in the ray $\{(0, ct, t) : t \geq 0\}$.

[32] We thank Yossi Feinberg for suggesting the Lyapunov Convexity Theorem, which simplified an earlier proof.

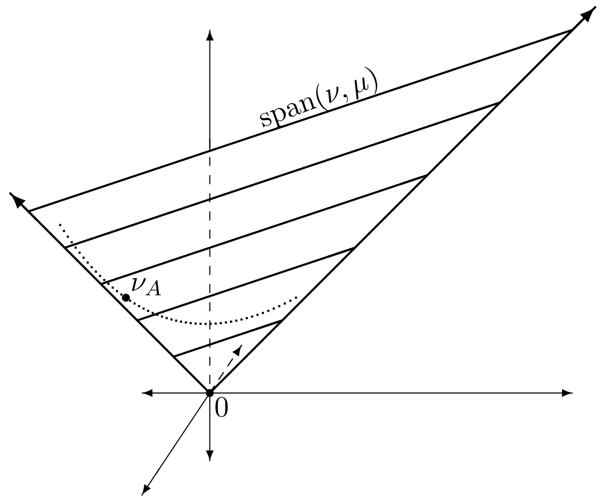

Figure 2: Lemma 18

Since ν_A has strictly positive components (namely the strict upper contour set of A), the set

$$\{x \in [\nu_A, \nu_B, \nu_C](\mathcal{Z}) : x_1 > 0\}$$

is nonempty. Therefore, we can select vectors $x^0 \in [\nu_A, \nu_B, \nu_C](\mathcal{Z})$ such that $x_1^0 = 0$ and $x^1 \in [\nu_A, \nu_B, \nu_C](\mathcal{Z})$ such that $x_1^1 > 0$. We show that the two vectors x^0, x^1 together span $[\nu_A, \nu_B, \nu_C](\mathcal{Z})$. Let $x^* \in [\nu_A, \nu_B, \nu_C](\mathcal{Z})$. We proceed by cases.

Case 1: Suppose $x_1^* = 0$. Since $\{x \in \mathbb{R}^3 : x_1 = 0\} \cap [\nu_A, \nu_B, \nu_C](\mathcal{Z})$ is a ray, $x^* = cx^0$.

Case 2: Suppose $x_1^* < 0$. Since $x_1^1 > 0$ is nonempty, we can find a convex combination $x' = \alpha x^* + (1 - \alpha)x^1$ with $x'_1 = 0$. The vector x' is in the range of the vector-valued measure $[\nu_A, \nu_B, \nu_C]$, because it is a convex combination of two elements. But, applying Case 1 to x', we conclude x' is spanned by x^0. x^* is spanned by x' and x^1, while x' is a multiple of x^0. So x^* is spanned by x^0 and x^1.

Case 3: Suppose $x_1^* > 0$. Apply Case 2 to $-x_1^*$.

Hence, there exist constants α, β such that $\nu_A = \alpha\nu_B + \beta\nu_C$. This suffices to show the entire space $\{\nu_A : A \in \mathcal{Z}\}$ can be spanned by any two of its measures, since the selection of A, B, C in the proof was arbitrary.

We finally show this span contains a strictly positive measure. Rescale each ν_A so $\|\nu_A\| = 1$. The space of probability measures $\Delta(\mathcal{Z})$ is a closed subset of $ca(\mathcal{Z})$, while ν_A approaches $\Delta(\mathcal{Z})$ as A approaches the $\succsim$–minimal element a_* in the order topology of $\succsim$, which is coarser than the Euclidean topology by continuity. Since any subspace of $ca(\mathcal{Z})$ is closed, the span contains a probability measure, which we can use as μ.

The lemma is illustrated in Figure 2. The three-dimensional space is supposed to capture the infinite-dimensional space $ca(\mathcal{Z})$. Drawn with heavier lines, the plane cutting through the origin is the span of ν and μ. The measures ν_A, drawn as the dotted curve, live inside that span. A particular ν_A is labeled inside the curve. $\square$

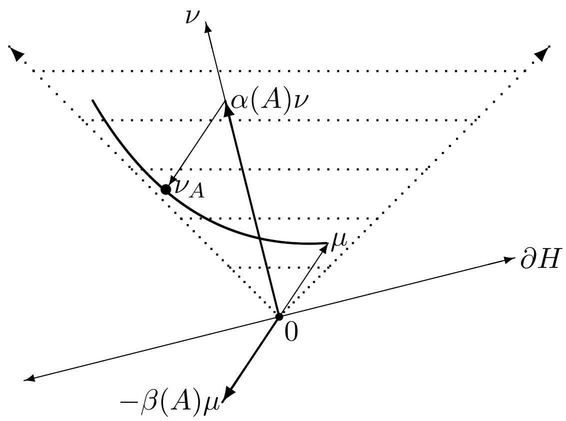

Figure 3: Constructions in the proof of Lemma 20

Lemma 19. *The set $\{\alpha\nu_A : A \in \mathcal{Z}, \alpha > 0\}$ is a convex positive cone.*

Proof. Take ν_A, ν_B. If $A \sim B$, then $\nu_A = \alpha\nu_B$ for some constant $\alpha > 0$ and any convex combination of ν_A and ν_B is immediately in $\{\alpha\nu_A : A \in \mathcal{Z}, \alpha > 0\}$. So assume that $A \succ B$. Then ν_A and ν_B are linearly independent and can serve as a basis for $\text{span}(\nu, \mu)$. Let $\nu_\alpha = \alpha\nu(A) + (1-\alpha)\nu_B$.

There exists some C such that $\nu_\alpha(C) = 0$. Since $\nu_A \geq 0, \nu_B(C) > 0$ whenever $C \succsim A$ and $\nu_B(C) \leq 0, \nu_A(C) < 0$ whenever $C \precsim B$, it must be the case that $A \succ C \succ B$. Then ν_B, ν_C are linearly independent and can serve as a basis for $\text{span}(\nu, \mu)$. Pick $\gamma < \frac{\nu_C(A)}{\nu_B(A)}$. Notice $\gamma > 0$ since $\nu_C(A), \nu_B(A) > 0$. Since ν_B, ν_C are linearly independent, $\gamma\nu_C, \nu_B$ are also linearly independent and can serve as a basis for $\text{span}(\nu, \mu)$. Let (q, r) denote the coordinates for ν_A with respect to this basis.

Since $\nu_A(A) = 0$ we have $q(\nu_B(A)) + r(\gamma\nu_C(A)) = \nu_A(A) = 0$. By our selection of γ, we have $q < 0 < r$. Let $c = 1/r\gamma > 0$ and $d = -q/r\gamma > 0$. Then $\nu_C = c\nu_A + d\nu_B$. Let $\beta = c/(c+d) > 0$. Simple algebra verifies that $\nu_\alpha = \beta\nu_C$. $\square$

Lemma 20. *There exists some ν^* in span(ν, μ) such that $\frac{\nu^*(A)}{\mu(A)}$ is a utility representation.*

Proof. For any $\gamma \in \text{span}(\mu, \nu)$, let $\alpha(\gamma), \beta(\gamma)$ solve $\alpha(\gamma)\nu + \beta(\gamma)\mu = \gamma$. Take any $B \succsim C$. If $B \sim C$, then $\nu_B = \nu_C$. If $B \succ C$, there exists some D with $B \succ D \succ C$, implying $\nu_B(D) > 0$ and $\nu_C(D) < 0$. Since $\nu_B(A) > 0$ and $\nu_C(A) > 0$, there exists no $\alpha > 0$ such that $\nu_B = -\alpha\nu_C$. In either case, it is impossible for $\nu_B = -\alpha\nu_C$ for any $\alpha > 0$. As B, C are arbitrary and $\{\gamma\nu_S : S \in \mathcal{Z}, \gamma > 0\}$ is a convex positive cone by Lemma 19, there exists some half space H in $ca(\mathcal{Z})$ such that $\{\nu_S : S \in \mathcal{Z}\} \subseteq H$. Let $H^* = H \cap \text{span}(\mu, \nu)$. This H^* is a two-dimensional half space of $\text{span}(\mu, \nu)$, so H^* is defined by $\{\gamma \in \text{span}(\mu, \nu) : a\alpha(\gamma) + b\beta(\gamma) + c = 0\}$ for some constants a, b, c.

Furthermore, for any $S \in \mathcal{Z}$, there exists $T \succ S$, for which $\nu_S(T) > 0$ and $\mu(T) \geq 0$. Therefore, there exists no $\alpha > 0$ or ν_S such that $\nu_S = -\alpha\mu$. Thus, we may proceed without loss of generality by assuming $a\alpha(\mu) + b\beta(\mu) + c = 0$. Pick $\nu^* \in \text{span}(\mu, \nu)$ such that $a\alpha(\nu^*) + b\beta(\nu^*) > 0$. Obviously, μ and ν^* are linearly independent. Therefore, we can find $\alpha^*(A), \beta^*(A)$ such that $\alpha^*(A)\nu^* + \beta^*(A)\mu = \nu_A$. By the selection of ν^*, we must have $\alpha^*(\nu_A) > 0$ to satisfy the inequality $\alpha(\nu_A) + \beta(\nu_B) + c \geq 0$.

Let $\alpha(A) = \alpha^*(\nu_A)$ and $\beta(A) = -\beta^*(\nu_A)$. By construction,

$$\alpha(A)\nu^*(A) - \beta(A)\mu(A) = 0 = \nu_A(A).$$

Then

$$\frac{\beta(A)}{\alpha(A)} = \frac{\nu^*(A)}{\mu(A)},$$

so it suffices to show that the left hand side is a representation for $\succsim$. Take $A \succsim B$. By construction,

$$\nu_B(A) \geq 0 = \nu_A(A).$$

Also, our selection of $\alpha(A)$ and $\beta(A)$ implies

$$\begin{aligned}
\nu(A) - \frac{\beta(A)}{\alpha(A)}\mu(A) &= \frac{1}{\alpha(A)}\left[\alpha(A)\nu^*(A) - \beta(A)\mu(A)\right] \\
&= \frac{\nu_B(A)}{\alpha(A)} \\
&\geq 0,
\end{aligned}$$

the last inequality following from $\alpha(A) > 0$. Similarly,

$$\nu^*(A) - \frac{\beta(B)}{\alpha(B)}\mu(A) \leq 0.$$

Together these two inequalities imply

$$\frac{\beta(A)}{\alpha(A)} \geq \frac{\beta(B)}{\alpha(B)}.$$

The argument when $A \succ B$ is identical, replacing weak inequalities with strict inequalities.

The proof is illustrated in Figure **??**. The two-dimensional figure shows the span of ν and μ, laying the plane in Figure 2 flat against the page. The line ∂H is the border that defines the half space H in which all of the ν_A's live; ν is normal to that border. The coefficients $\alpha(A)$ and $\beta(A)$ on ν and μ are shown for a particular set A; the coefficients are defined by $\nu_A = \alpha(A)\mu - \beta(A)\nu$. □

Proof of Theorem. Invoking the Radon–Nikodym Theorem, we can rewrite $\nu^*(A)$ of Lemma 20 as $\int_A u \, d\mu$, where u is the Radon–Nikodym derivative of ν with respect to μ. The absolute continuity condition on these measures holds because $\mu(A) > 0$ for any $A \in \mathcal{Z}$ by disjoint set betweenness. The continuity of u is guaranteed by the downward Hausdorff continuity axiom and Lemma 6. the full support of μ follows from the strict part of disjoint set betweenness. This proves the sufficiency of the axioms for the representation on the regular sets. The representation for singletons follows immediately from Lemma 6. □

References

Ahn, D. S. (forthcoming): "Hierarchies of Ambiguous Beliefs," *Journal of Economic Theory.*

Aliprantis, C. D., and K. C. Border (1999): *Infinite Dimensional Analysis*. Springer, New York, 2nd edn.

Anscombe, F. J., and R. J. Aumann (1963): "A Definition of Subjective Probability," *Annals of Mathematical Statistics*, 34, 199–205.

Arrow, K. J., and L. Hurwicz (1972): "An Optimality Criterion for Decision-Making under Ignorance," in *Uncertainty and Expectations in Economics: Essays in Honour of G. L. S. Shackle*, ed. by C. F. Carter, and J. L. Ford, pp. 1–11. Blackwell, Oxford.

Birkhoff, G., and J. von Neumann (1936): "The Logic of Quantum Mechanics," *Annals of Mathematics*, 37, 823–843.

Blume, L., A. Brandenburger, and E. Dekel (1991): "Lexicographic Probabilities and Choice Under Uncertainty," *Econometrica*, 59, 61–79.

Bolker, E. D. (1966): "Functions Resembling Quotients of Measures," *Transactions of the American Mathematial Society*, 124, 292–312.

——— (1967): "A Simultaneous Axiomatization of Utility and Subjective Probability," *Philosophy of Science*, 34, 333–340.

Broome, J. (1990): "Bolker–Jeffrey Expected Utility Theory and Axiomatic Utilitarianism," *Review of Economic Studies*, 57, 477–502.

Dekel, E. (1986): "An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom," *Journal of Economic Theory*, 40, 304–318.

Dekel, E., B. L. Lipman, and A. Rustichini (2001): "Representing Preferences with a Unique Subjective State Space," *Econometrica*, 69, 891–934.

Dunford, N., and J. T. Schwartz (1957): *Linear Operators: Part I.* Interscience, New York.

Ellsberg, D. (1961): "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75, 643–669.

Epstein, L. G. (1999): "A Definition of Uncertainty Aversion," *Review of Economic Studies*, 66, 579–608.

Epstein, L. G., M. Marinacci, and K. Seo (2007): "Coarse Contingencies," Working paper, University of Rochester and Università di Torino.

Epstein, L. G., and J. Zhang (2001): "Subjective Probabilities on Subjectively Unambiguous Events," *Econometrica*, 69(2), 265–306.

Ergin, H., and F. Gul (2004): "A Subjective Theory of Compound Lotteries," Working paper, Princeton University.

Gajdos, T., T. Hayashi, J.-M. Tallon, and J.-C. Vergnaud (2006): "Attitude toward Imprecise Information," Working paper, Université Paris I Panthéon-Sorbonne and University of Texas, Austin.

Gajdos, T., J.-M. Tallon, and J.-C. Vergnaud (2004): "Decision Making with Imprecise Probabilistic Information," *Journal of Mathematical Economics*, 40, 647–681.

Ghirardato, P. (2001): "Coping with Ignorance: Unforeseen Contingencies and Non-Additive Uncertainty," *Economic Theory*, 17, 247–276.

Ghirardato, P., F. Maccheroni, and M. Marinacci (2004): "Differentiating Ambiguity and Ambiguity Attitude," *Journal of Economic Theory*, 118, 133–173.

Ghirardato, P., and M. Marinacci (2002): "Ambiguity Made Precise: A Comparative Foundation," *Journal of Economic Theory*, 102, 251–289.

Gilboa, I., and D. Schmeidler (1989): "Maxmin Expected Utility with a Non-Unique Prior," *Journal of Mathematical Economics*, 18, 141–153.

Giraud, R. (2006): "Objective Imprecise Probabalistic Information, Second Order Beliefs and Ambiguity Aversion: An Axiomatization," Working paper, University of Franche-Comté.

Gul, F., and W. Pesendorfer (2001): "Temptation and Self–Control," *Econometrica*, 69, 1403–1435.

Halevy, Y. (2007): "Ellsberg Revisited: An Experimental Study," *Econometrica*, 75, 503–536.

Halevy, Y., and V. Feltkamp (2005): "A Bayesian Approach to Uncertainty Aversion," *Review of Economic Studies*, 72, 449–466.

Halmos, P. R. (1974): *Measure Theory.* Springer, New York, Reprint.

Hayashi, T., and R. Wada (2007): "Choice with Imprecise Information: An Experimental Approach," Working paper, University of Texas, Austin and Keiai University.

Jaffray, J.-Y. (1989): "Linear Utility Theory for Belief Functions," *Operations Research Letters*, 8, 107–112.

Jaffray, J.-Y., and M. Jeleva (2004): "How to Deal with Partially Analyzed Acts? A Proposal," Working paper, University Paris 6 and University of Nantes.

Jeffrey, R. C. (1965): *The Logic of Decision*. McGraw-Hill, New York.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, 73, 1849–1892.

Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky (1971): *Foundations of Measurement*, vol. 1. Academic Press, New York.

Kreps, D. M. (1979): "A Representation Theorem for 'Preference for Flexibility'," *Econometrica*, 47, 565–576.

Machina, M. J., and D. Schmeidler (1992): "A More Robust Definition of Subjective Probability," *Econometrica*, 60, 745–780.

Mukerji, S. (1997): "Understanding the Nonadditive Probability Decision Model," *Economic Theory*, 7, 23–46.

Nau, R. F. (2006): "Uncertainty Aversion with Second-Order Utilities and Probabilities," *Management Science*, 52, 136–145.

Nehring, K. (1999): "Preference for Flexibility in a Savage Framework," *Econometrica*, 67, 101–119.

——— (2001): "Ambiguity in the Context of Probabilistic Beliefs," Working paper, University of California, Davis.

Olszewski, W. B. (2007): "Preferences over Sets of Lotteries," *Review of Economic Studies*, 74, 567–595.

Savage, L. J. (1954): *The Foundations of Statistics*. Wiley, New York.

Segal, U. (1987): "The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach," *International Economic Review*, 58, 349–377.

——— (1990): "Two-Stage Lotteries without the Reduction Axiom," *Econometrica*, 58, 349–377.

Seo, K. (2006): "Ambiguity and Second-Order Belief," Working paper, University of Rochester.

Siniscalchi, M. (2006): "A Behavioral Characterization of Plausible Priors," *Journal of Economic Theory*, 128, 91–135.

Stinchcombe, M. B. (2003): "Choice and Games with Ambiguity as Sets of Probabilities," Working paper, University of Texas, Austin.

Suppes, P. (1966): "The Probabilistic Argument for a Non-Classical Logic of Quantum Mechanics," *Philosophy of Science*, 33, 14–21.

Viero, M.-L. (2006): "Exactly What Happens after the Anscombe-Aumann Race? Representing Preferences in Vague Environments," Working paper, Queen's University.

von Neumann, J. (1955): *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, Princeton, translated by R. T. Beyer.

von Neumann, J., and O. Morgenstern (1944): *Theory of Games and Economic Behavior*. Princeton University Press, Princeton.

Wang, T. (2003): "A Class of Multi–Prior Preferences," Working paper, University of British Columbia.

Zhang, J. (1999): "Qualitative Probabilities on λ-Systems," *Mathematical Social Sciences*, 38, 11–20.